ALGEBRA LINEAL

Héctor Jairo Martínez R.
Ana María Sanabria R.

SEGUNDO SEMESTRE 2008

Índice general

Capítulo 1

SISTEMAS DE ECUACIONES LINEALES

1.1. Introducción

Iniciamos este texto con la presentación de los Sistemas de Ecuaciones Lineales, ya que son una herramienta importante para el análisis y el manejo de los conceptos fundamentales del Álgebra Lineal que abordaremos en los capítulos posteriores.

Aprovechando el concepto intuitivo que sobre Sistemas de Ecuaciones tiene el lector, motivaremos los conceptos de matriz y de vector de R^n como una forma de simplificar la escritura, análisis y solución de los Sistemas de Ecuaciones Lineales. Posteriormente, presentaremos una técnica para resolverlos, haciendo énfasis en la determinación de si tiene o no solución y cuando la tiene, si ésta es única o no, sin necesidad de calcularla. La técnica que utilizaremos, por su fácil comprensión, es la clásica eliminación de Gauss seguida de una sustitución regresiva, basada en las operaciones elementales, que muy intuitivamente se pueden ver como operaciones entre ecuaciones.

El planteamiento de Sistemas de Ecuaciones Lineales es un problema básico e importante de la Matemática, al igual que el diseño de métodos eficientes de solución de ellos. Ambos problemas, muy relacionados con el desarrollo del Álgebra Lineal, son el objeto de estudios de cursos como Modelamiento Matemático, el primero, y Análisis Numérico, el segundo; por lo tanto, no serán tratados a profundidad en este texto.

1.2. Conceptos Básicos

Para empezar, recordemos qué es una *ecuación* y el significado de *solución* de la misma.

Definición 1 [*Ecuación*]. Una ecuación con incógnitas o variables $x_1, x_2, \cdots, x_n$ es una proposición abierta[1] que establece la igualdad entre dos expresiones matemáticas que explícitamente involucran todas, algunas o ninguna de estas variables.

Ejemplo 1.

1. La proposición
$$3x_1^2 - x_2 = 4x_1 + x_3 \tag{1.1}$$
es una ecuación con variables x_1, x_2, x_3 .

2. La proposición
$$x - 2y + 3z + 1 = s + r - 2, \tag{1.2}$$
considerando a s y r como constantes reales(parámetros), es una ecuación con variables x, y, z.

[1] Una proposición abierta o funcional es aquella que establece su valor de verdad dependiendo de los valores que tomen las variables

3. La proposición
$$5-3=2 \tag{1.3}$$
podemos considerarla como una ecuación con cualquier número de variables. Por ejemplo, podemos pensarla con variables x, y, z, w $(0x+0y^2+5+0z-3=2+0w)$ o podemos pensarla con variables x_1, x_2 $(5+0x_1-3=0x_2+2)$.

4. Igualmente, podemos considerar la proposición
$$7+3=5 \tag{1.4}$$
como una ecuación con variables x_1, x_2, x_3 $(7+0x_1^3+3=5+0x_2+0x_3)$. □

Definición 2 [*Solución de una ecuación*]. Una solución de una ecuación con variables $x_1, x_2, \cdots, x_n$ es una n-upla tal que, al sustituir cada una de las variables de la ecuación por las componentes respectivas de la n-upla, obtenemos una identidad (proposición universalmente verdadera). Al conjunto formado por todas las soluciones de una ecuación lo llamaremos *conjunto solución.*

Ejemplo 2.

1. La tripla $(1,-2,1)$ es una solución de la Ecuación (1.1), ya que al sustituir x_1 por 1, x_2 por -2 y x_3 por 1, obtenemos 5=5. De la misma forma, la tripla $(1,0,1)$ no es solución de esta ecuación, ya que al sustituir x_1 por 1, x_2 por 0 y x_3 por 1, obtenemos $3=5$.

2. La tripla $(3s+3r, s+r, -1)$ es solución de la Ecuación (1.2), para cualquier valor de s y r, ya que al sustituir x por $(3s+3r)$, y por $(s+r)$ y z por -1, obtenemos $s+r-2=s+r-2$.

3. La Ecuación (1.3) siempre tiene solución, independientemente de los valores de las variables, obtendremos la identidad $2=2$. □

4. La Ecuación (1.4) no tiene solución, ya que independientemente de los valores por los que sustituyamos a x_1, x_2, x_3, obtendremos la proposición falsa $10=5$; es decir, el **conjunto solución** de esta ecuación es ϕ, el conjunto vacío.

En particular, en este capítulo, nos ocuparemos de un tipo especial de ecuaciones, las ecuaciones lineales.

Definición 3 [*Ecuación lineal*]. Una ecuación que se puede escribir de la forma
$$a_1x_1+a_2x_2+\cdots+a_nx_n=b,$$
es una *ecuación lineal* con variables $x_1, \cdots, x_n$, coeficientes $a_1, \cdots, a_n$ y término independiente b reales[2]. Al primer coeficiente diferente de cero lo llamamos *pivote de la ecuación* y a la variable correspondiente, *variable pivotal de la ecuación*. Cuando $b=0$, a la ecuación la denominamos *ecuación lineal homogénea.*

Observemos que la variable pivotal se puede despejar en función de las demás variables. Notemos sin embargo que la variable pivotal no siempre es la única variable que se puede despejar en término de las demás. Esta observación será de mucha importancia para la solución de un sistema de ecuaciones lineales.

Dada una ecuación lineal con término b arbitrario, definimos la *ecuación lineal homogénea asociada* a ésta, como la ecuación con las mismas variables, de coeficientes iguales y término independiente $b=0$.

Ejemplo 3. La ecuación
$$2x-y=7 \tag{1.5}$$
es una ecuación lineal con variables x, y, donde el coeficiente de x es 2, el coeficiente de y es -1, el término independiente es 7, el pivote es 2, la variable pivotal es x y la ecuación lineal homogénea asociada es
$$2x-y=0.$$

[2]Cada término de la ecuación contiene a lo más una sola variable con exponente 1

Se puede verificar que (5, 3), (0, -7) y en general todas las duplas de la forma $\left(\frac{7+t}{2},\ t\right)$, donde t es un número real, son soluciones de la ecuación original, mientras que (0, 0), (4, 8) y en general todas las duplas de la forma $\left(\frac{t}{2},\ t\right)$, donde t es un número real, son soluciones de la ecuación homogénea asociada; es decir, los conjuntos solución de la ecuación (1.5) y la ecuación homogénea asociada son

$$\left\{\left(\frac{7+t}{2},\ t\right),\, t \in R\right\} \quad \text{y} \quad \left\{\left(\frac{t}{2},\ t\right),\, t \in R\right\},$$

respectivamente. □

Observemos que el conjunto solución de la ecuación (1.5) también puede ser escrito como

$$\{(u,\ 2u-7)\,,\, u \in R\}\,.$$

En efecto, al reemplazar x por u y a y por $2u-7$ en (1.5), obtenemos $2u-(2u-7) = 2u-2u+7 = 7$.

Ahora, si tenemos un número finito de ecuaciones lineales con las mismas variables, podemos agruparlas en lo que llamamos un *sistema de ecuaciones lineales*.

Definición 4 [*Sistema de Ecuaciones Lineales*]. Un *sistema de m ecuaciones lineales con n variables* $x_1, \cdots, x_n$ es un conjunto de m ecuaciones lineales de la forma

$$\begin{array}{rcl} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n &=& b_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n &=& b_2 \\ \vdots & & \vdots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n &=& b_m. \end{array} \tag{1.6}$$

El número α_{ij} es el coeficiente de la variable x_j en la ecuación i y b_i es el término independiente de la ecuación i. Cuando **todos** los términos independientes b_i son 0, el sistema lo llamamos *homogéneo*.

Dado un sistema de ecuaciones con términos independientes b_i, se define *el sistema de ecuaciones lineales homogéneo asociado* a éste, como el sistema de ecuaciones lineales con las mismas variables, de coeficientes iguales y términos independientes $b_i = 0$.

Ejemplo 4. El siguiente conjunto de ecuaciones

$$\begin{array}{rcrcrcr} x_1 &+& 2x_2 & & &=& -3 \\ 2x_1 &+& 3x_2 &-& 2x_3 &=& -10 \end{array}$$

es un sistema de 2 ecuaciones lineales con variables x_1, x_2, x_3, donde los coeficientes respectivos de x_1, x_2 y x_3, en la primera ecuación, son 1, 2 y 0; y, en la segunda ecuación, son 2, 3 y -2. Los términos independientes de la primera y segunda ecuación son -3 y -10, respectivamente. El sistema

$$\begin{array}{rcrcrcr} x_1 &+& 2x_2 & & &=& 0 \\ 2x_1 &+& 3x_2 &-& 2x_3 &=& 0 \end{array}$$

es su sistema homogéneo asociado. □

Definición 5 [*Solución de un sistema de ecuaciones lineales*]. Una *solución de un sistema de ecuaciones lineales* con variables $x_1, x_2, \cdots, x_n$ es una n-upla que es solución de todas y cada una de las ecuaciones del sistema. Al conjunto formado por todas las soluciones de un sistema lo llamamos *conjunto solución.*

Ejemplo 5. Podemos verificar que las triplas $(-15, 6, -1)$ y $(4, -2, 1)$ son, respectivamente, soluciones de los sistemas del ejemplo anterior, ya que al hacer los debidos reemplazos, obtenemos

$$\begin{array}{rcrcrcr} 1\cdot(-15) &+& 2\cdot 6 &+& 0\cdot(-1) &=& -3 \\ 2\cdot(-15) &+& 3\cdot 6 &+& -2\cdot(-1) &=& -10 \end{array} \qquad \text{es decir,} \qquad \begin{array}{rcr} -3 &=& -3 \\ -10 &=& -10 \end{array}$$

y

$$\begin{array}{rcrcrcr} 1\cdot 4 & + & 2\cdot(-2) & & & = & 0 \\ 2\cdot 4 & + & 3\cdot(-2) & + & -2\cdot 1 & = & 0 \end{array} \quad \text{es decir,} \quad \begin{array}{rcr} 0 & = & 0 \\ 0 & = & 0 \end{array}.$$

De la misma manera, podemos ver que $(1,-2,4)$ no es solución del sistema original ya que, aunque es solución de la primera ecuación $(1\cdot1+2\cdot(-2)=-3)$, no es solución de la segunda ecuación $(2\cdot1+3\cdot(-2)-2\cdot4 \neq -10)$.□

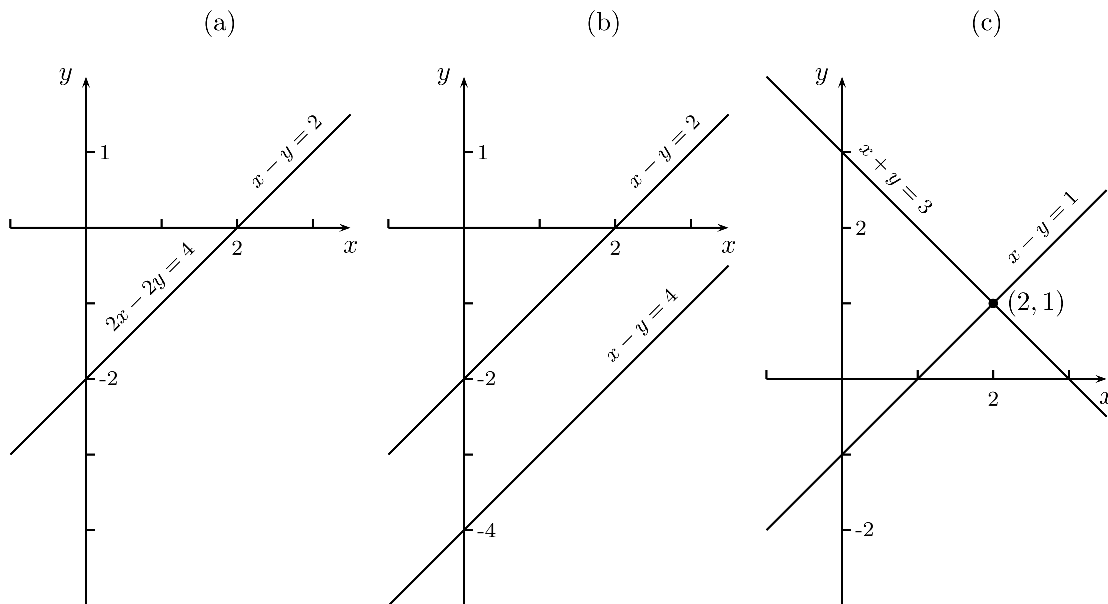
Observemos que todo sistema de ecuaciones lineales homogéneo siempre tiene al menos una solución: todas las variables iguales a cero (*solución trivial*).

Desde el punto de vista geométrico, al *conjunto solución* de un sistema de ecuaciones lineales lo podemos interpretar como la intersección de todos los lugares geométricos (rectas, planos o hiperplanos) representados por cada una de las ecuaciones lineales. Por simplicidad, aunque lo ampliaremos más adelante, ilustraremos lo anterior en el caso de dos ecuaciones con dos variables.

Ejemplo 6. Los siguientes tres conjuntos de ecuaciones son sistemas de 2 ecuaciones con variables x, y. Veremos que el número de soluciones de cada uno de ellos es completamente diferente.

(a) $\begin{array}{rcrcr} x & - & y & = & 2 \\ 2x & - & 2y & = & 4 \end{array}$ (b) $\begin{array}{rcrcr} x & - & y & = & 2 \\ x & - & y & = & 4 \end{array}$ (c) $\begin{array}{rcrcr} x & - & y & = & 1 \\ x & + & y & = & 3 \end{array}$

Observemos que, en el sistema (a), la segunda ecuación es la primera multiplicada por 2, de tal manera que las soluciones de las dos ecuaciones son las mismas. De otro lado, es fácil verificar que toda dupla de la forma $(r+2,r)$, donde r es cualquier número real, es solución de este sistema y que por tanto, el sistema tiene *infinitas soluciones*. Geométricamente, podemos ver que cada una de las ecuaciones del sistema corresponde a la ecuación de una misma recta en el plano, así que las coordenadas de todos los puntos de dicha recta son las soluciones del sistema.



Respecto al sistema (b), podemos ver que *no tiene solución* ya que no hay dos números que restados den 2 y 4, simultáneamente. Geométricamente, se puede verificar que las ecuaciones del sistema corresponden a las ecuaciones de dos rectas paralelas, así que no hay un punto común a ellas y por tanto, no hay una dupla que sea solución del sistema.

Y, respecto al sistema (c), podemos ver que $(2,1)$ es solución del sistema y que además es la *única solución*. Geométricamente, las ecuaciones del sistema corresponden a las ecuaciones de dos rectas no paralelas en el plano; por tanto, se cortan en un sólo punto: $(2,1)$. □

En este ejemplo, se ilustran las únicas tres posibilidades del número de soluciones de un sistema de ecuaciones lineales: infinitas(a), ninguna(b) y sólo una(c) , como lo veremos más adelante. Por lo pronto, diferenciaremos muy bien aquellos sistemas que tienen al menos una solución de aquellos que no tienen solución.

Definición 6 [*Sistema de ecuaciones consistente*]. A un sistema de ecuaciones lo denominamos *consistente* si tiene al menos una solución. En caso contrario, al sistema lo denominamos *inconsistente*.

En el Ejemplo 6, tenemos que los sistemas (a) y (c) son consistentes, mientras que el sistema (b) es inconsistente. Además, podemos afirmar que todo sistema de ecuaciones lineales homogéneo es consistente (Por qué?) y que el conjunto solución de un sistema inconsistente es el conjunto vacío

Para determinar el número de soluciones y/o encontrar las soluciones de un sistema, recurriremos, como lo ilustraremos más adelante, a sistemas que tienen las mismas soluciones y que son más fácil de analizar y de resolver.

Definición 7 [*Sistemas de ecuaciones equivalentes*]. Decimos que dos sistemas son *equivalentes*, si tienen el mismo conjunto de soluciones.

Ejemplo 7. Los siguientes sistemas son equivalentes.

$$(a)\ \begin{array}{ccccc} x & - & y & = & 1 \\ x & + & y & = & 3 \end{array} \qquad\qquad (b)\ \begin{array}{ccccc} x & - & y & = & 1 \\ & & y & = & 1 \end{array}$$

ya que ambos tienen como única solución (2, 1). □

De otro lado, es fácil ver que los sistemas que tienen el "patrón escalonado" del sistema (b) del Ejemplo 7, son fáciles de resolver. En efecto, si comenzamos resolviendo la última ecuación y luego sustituimos el valor encontrado en la ecuación anterior y resolvemos la ecuación resultante, obtenemos la solución del sistema; es decir,

$$\begin{aligned} y &= 1, \\ x-1 &= 1 \text{ de donde, } x=2. \end{aligned}$$

El procedimiento empleado para resolver este último sistema se conoce como *sustitución hacia atrás* y la idea básica para resolver un sistema cualquiera consiste en transformar el sistema dado en otro equivalente que tenga el "patrón escalonado", el cual como ya dijimos, será más fácil de resolver. Esta idea es lo que en el bachillerato llamábamos método de reducción o eliminación. Veamos un ejemplo de solución mediante esta idea básica, la cual formalizaremos en la Sección 1.4.

Ejemplo 8. Resolvamos el sistema de ecuaciones lineales

$$\begin{array}{ccccccc} x & - & y & - & z & = & 2 \\ 3x & - & 3y & + & 2z & = & 16 \\ 2x & - & y & + & z & = & 9. \end{array} \tag{1.7}$$

Restándole, a la segunda ecuación, 3 veces la primera ecuación para eliminar la variable x de la segunda ecuación, obtenemos el siguiente sistema:

$$\begin{array}{ccccccc} x & - & y & - & z & = & 2 \\ & & & & 5z & = & 10 \\ 2x & - & y & + & z & = & 9. \end{array} \tag{1.8}$$

Restándole, a la tercera ecuación, 2 veces la primera ecuación para eliminar la variable x de la tercera ecuación, obtenemos el siguiente sistema:

$$\begin{array}{ccccccc} x & - & y & - & z & = & 2 \\ & & & & 5z & = & 10 \\ & & y & + & 3z & = & 5. \end{array} \tag{1.9}$$

Intercambiando la segunda y la tercera ecuación, obtenemos el siguiente sistema:

$$\begin{array}{rcrcrcr} x & - & y & - & z & = & 2 \\ & & y & + & 3z & = & 5 \\ & & & & 5z & = & 10, \end{array} \tag{1.10}$$

el cual tiene un "patrón escalonado", que hace fácil encontrar la solución mediante el método de sustitución hacia atrás: De la última ecuación, $z = 2$. Sustituyendo el valor de z en la ecuación inmediatamente anterior, es fácil ver que $y = -1$; y sustituyendo los valores de y y z en la ecuación inmediatamente anterior a ésta última, tenemos que $x = 3$. □

Los procedimientos que hicimos para pasar de un sistema a otro, nos permiten ver que las soluciones del sistema inicial son también solución de cada uno de los otros sistemas. Puesto que los pasos que efectuamos son reversibles, podríamos iniciar el proceso en el sistema final y pasando por cada uno de los otros sistemas llegar al sistema inicial. Así, cualquier solución del sistema final es también solución de los otros sistemas y en particular, del sistema inicial.

Observemos que, en el sistema (1.10), el cual tiene el "patrón escalonado" la variable pivotal de cada ecuación es diferente, lo cual permite obtener la solución del sistema mediante la sustitución hacia atrás.

Volviendo a los procedimientos para pasar de un sistema a otro equivalente, debemos anotar que sólo utilizamos dos tipos de operaciones entre ecuaciones de los tres tipos que definimos a continuación.

Definición 8 [*Operaciones elementales entre ecuaciones*]. Dado un sistema de ecuaciones lineales, llamamos *operaciones elementales entre ecuaciones* a cada uno de las siguientes procedimientos:
Escalamiento. Reemplazar la ecuación i, E_i, por un múltiplo de ésta, cE_i, $c \neq 0$: $cE_i \longrightarrow E_i$.
Eliminación. Reemplazar la ecuación i, E_i, por la suma de ésta con un múltiplo de otra, cE_j: $E_i + cE_j \longrightarrow E_i$.
Permutación. Intercambiar las ecuaciones i y j, E_i y E_j: $E_i \longleftrightarrow E_j$.

Con esta notación, tenemos que, en el Ejemplo 8, las operaciones elementales que efectuamos para pasar del sistema (1.7) al sistema (1.8), del sistema (1.8) al sistema (1.9) y del sistema (1.9) al sistema (1.10) son $E_2 - 3E_1 \longrightarrow E_2$, $E_3 - 2E_1 \longrightarrow E_3$, y $E_2 \longleftrightarrow E_3$, respectivamente. Así, podemos observar que después de efectuar sucesivamente estas operaciones elementales a un sistema, se llega a otro equivalente

1.3. Representación Matricial de un Sistema de Ecuaciones Lineales

Observemos que los sistemas de ecuaciones lineales se caracterizan por el valor de los coeficientes de sus variables y los términos independientes, de manera que, para empezar, podemos ahorrarnos la escritura repetitiva de las variables. Para tal efecto, llamamos *matriz de coeficientes del sistema* (o simplemente, *matriz del sistema*) al arreglo rectangular de números formado por los coeficientes de las variables, de tal forma que cada *fila* corresponda a una ecuación y cada *columna* a una variable, y *matriz aumentada del sistema*, al arreglo rectangular de números formado por la matriz del sistema y una columna adicional conformada por los términos independientes. De esta forma, la matriz de un sistema con m ecuaciones y n variables tendrá m filas y n columnas, lo cual denotamos diciendo que el tamaño de la matriz $m \times n$. Así, la matriz del sistema y la matriz aumentada del sistema del Ejemplo 4 son, respectivamente, las matrices de tamaño 2×3 y 2×4

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & -2 \end{pmatrix} \quad \text{y} \quad \left(\begin{array}{rrr|r} 1 & 2 & 0 & -3 \\ 2 & 3 & -2 & -10 \end{array}\right),$$

y en general, la matriz de coeficientes y la matriz aumentada del sistema de ecuaciones lineales (1.6) son

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \quad \text{y} \quad \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & b_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} & b_m \end{array}\right)^{3}.$$

Como la matriz aumentada de un sistema es una representación abreviada de éste, diremos que el pivote de una fila de una matriz es el primer elemento de la fila, de izquierda a derecha, que es diferente de cero[4] y las operaciones elementales entre ecuaciones se pueden interpretar como operaciones entre filas en esta matriz.

Definición 9 [*Operaciones elementales entre filas*]. Dada una matriz, llamaremos *operaciones elementales entre filas* a cada una de los siguientes procedimientos:
Escalamiento. Reemplazar la fila i, F_i, por un múltiplo de ésta, cF_i, $c \neq 0$: $cF_i \longrightarrow F_i$.
Eliminación. Reemplazar la fila i, F_i, por la suma de ésta con un múltiplo de otra, cF_j: $F_i + cF_j \longrightarrow F_i$.
Permutación. Intercambiar las filas i y j, F_i y F_j: $F_i \longleftrightarrow F_j$.

En el Ejemplo 8, los procedimientos efectuados para pasar de un sistema a otro, quedan representados por la siguientes matrices aumentadas y operaciones elementales entre filas:

$$\begin{array}{cc} & \left(\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right) \\ F_2 - 3F_1 \longrightarrow F_2 & \left(\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 2 & -1 & 1 & 9 \end{array}\right) \\ F_3 - 2F_1 \longrightarrow F_3 & \left(\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 0 & 1 & 3 & 5 \end{array}\right) \\ F_2 \longleftrightarrow F_3 & \left(\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right). \end{array} \tag{1.11}$$

Como estas matrices representan Sistemas de Ecuaciones Lineales equivalentes y fueron obtenidas mediante operaciones elementales entre filas, tenemos una motivación para la siguiente definición.

Definición 10 [*Matrices equivalentes*]. Decimos que dos matrices son equivalentes si al efectuar operaciones elementales entre filas a una de ellas, se obtiene la otra.

Ejemplo 9. Las matrices $\begin{pmatrix} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{pmatrix}$ y $\begin{pmatrix} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{pmatrix}$ son equivalentes, ya que como vimos en (1.11), al aplicar sucesivamente las operaciones elementales $F_2 - 3F_1 \longrightarrow F_2$, $F_3 - 2F_1 \longrightarrow F_3$ y $F_2 \longleftrightarrow F_3$ a la matriz $\begin{pmatrix} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{pmatrix}$, se obtiene la matriz $\begin{pmatrix} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{pmatrix}$. □

[3]Por notación, en la matriz aumentada, asi no aparezca la línea vertical, la última columna corresponde a los términos independientes.
[4]Observemos que aunque en una ecuación lineal el término independiente no puede ser pivote, con esta extensión del concepto de pivote, en una matriz aumentada, el término independiente si podría ser pivote.

1.4. Eliminación de Gauss

En esta sección, examinaremos con más detalle el método que utilizamos para resolver el sistema de ecuaciones lineales (1.7), el cual es un método sistemático y directo para resolver un sistema de ecuaciones lineales. La idea consiste en llevar la matriz aumentada del sistema dado a otra equivalente con "patrón escalonado", que corresponda a un sistema que pueda ser resuelto por medio de sustitución hacía atrás. Decimos que el método es directo porque, en un número finito de operaciones, se obtiene el conjunto solución del sistema.

Empezaremos formalizando lo que hemos llamado "patrón escalonado", definiendo lo que se conoce como matriz escalonada, ya que un Sistema de Ecuaciones Lineales con dicho patrón tiene como matriz aumentada una matriz escalonada.

Definición 11 [*Matriz escalonada*]. Decimos que una matriz es *escalonada* si posee las siguientes características:

1. Cualquier fila cuyas *componentes* sean todas cero, se encuentra en la parte inferior de la matriz.

2. En cada fila, cuyas *componentes* no sean todas cero, la primera componente diferente de cero (*pivote*), se encuentra en una columna que este a la derecha de cualquier otro *pivote* por encima de él (equivalentemente, podemos decir que cada pivote está a la izquierda de cualquier otro pivote por debajo de él).

Observemos que, en una matriz escalonada, ninguna columna tiene más de un pivote. Este hecho nos permite clasificar sus columnas como *pivotales* y *no pivotales*, según tengan o no pivote, respectivamente.

Ejemplo 10. Observemos que las siguientes matrices son matrices escalonadas.

$$\begin{pmatrix} \mathbf{1} & -2 & 1 & 1 \\ 0 & \mathbf{6} & -1 & -5 \\ 0 & 0 & \mathbf{-1/3} & 7/3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad , \quad \begin{pmatrix} \mathbf{3} & 1 & 0 & -2 & 0 \\ 0 & 0 & \mathbf{2} & 3 & 1 \\ 0 & 0 & 0 & 0 & \mathbf{-5} \end{pmatrix}$$

En efecto, en la primera matriz, sus pivotes son **1**, **6** y **-1/3** y sus columnas pivotales son la primera, la segunda y la tercera; y en la segunda matriz, sus pivotes son **3**, **2** y **-5** y sus columnas pivotales son la primera, la tercera y la quinta. □

Un método eficiente y de uso frecuente para reducir una matriz dada a una matriz escalonada equivalente es el conocido como método de eliminación de Gauss.

Método de Eliminación de Gauss
Objetivo: Escalonar una matriz; es decir, dada una matriz, encontrar una matriz escalonada equivalente a la matriz dada

1. Identifique la primera columna, de izquierda a derecha, que no sea de sólo ceros.

2. Si la primera componente de esta columna es un cero, intercambie la primera fila con una que tenga una componente no cero en esta columna. Esta componente no cero será el pivote de esta columna.

3. Aplique sucesivamente operaciones elementales adecuadas del Tipo Eliminación, para obtener ceros debajo del pivote de esta columna.

4. Repítale este procedimiento, comenzando con el Paso 1, al conjunto de filas de la matriz resultante del paso anterior que están por debajo de la fila donde está el pivote anterior, hasta agotar las filas o hasta que las filas restantes sean de sólo ceros.

Ejemplo 11. Aplicando el Método de Eliminación de Gauss, encontremos una matriz escalonada equivalente a la matriz

$$\begin{pmatrix} \mathbf{2} & -2 & 4 & 1 \\ 3 & 0 & 5 & -9/2 \\ 0 & 2 & -2/3 & -1 \\ 6 & -3 & 11 & -5 \end{pmatrix}.$$

Como la primera columna de izquierda a derecha que no es de sólo ceros es la primera columna de la matriz y su primer elemento no es cero, **2** es el pivote y, para obtener ceros debajo de él, realizamos las operaciones elementales entre filas indicadas a continuación

$$\begin{array}{ccc} F_2 - \frac{3}{2}F_1 & \longrightarrow & F_2 \\ F_4 - 3F_1 & \longrightarrow & F_4 \end{array} \qquad \begin{pmatrix} 2 & -2 & 4 & 1 \\ 0 & \mathbf{3} & -1 & -6 \\ 0 & 2 & -2/3 & -1 \\ 0 & 3 & -1 & -8 \end{pmatrix}.$$

Repitiendo el mismo procedimiento a las filas 3 y 4 de la matriz anterior y usando como segundo pivote el **3** de la fila 2, obtenemos

$$\begin{array}{ccc} F_3 - 2/3F_2 & \longrightarrow & F_3 \\ F_4 - F_2 & \longrightarrow & F_4 \end{array} \qquad \begin{pmatrix} 2 & -2 & 4 & 1 \\ 0 & 3 & -1 & -6 \\ 0 & 0 & 0 & \mathbf{3} \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Repitiendo nuevamente el mismo procedimiento a la fila 4 de la matriz anterior y usando como tercer pivote a **3** de la fila 3, obtenemos

$$F_4 + \tfrac{2}{3}F_3 \quad \longrightarrow \quad F_4 \qquad \begin{pmatrix} 2 & -2 & 4 & 1 \\ 0 & 3 & -1 & -6 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

la cual es una matriz escalonada equivalente a la matriz dada.

Notemos que la constante c de la operación elemental adecuada tipo eliminación es el opuesto del cociente entre la componente que se quiere anular y el pivote de la respectiva columna. Así, en el primer paso obtuvimos $-\frac{3}{2}$ y $-\frac{6}{2} = -3$; en el segundo, $-\frac{2}{3}$ y $-\frac{3}{3} = -1$ y, en el tercero, $-\frac{-2}{3} = \frac{2}{3}$. □

Observemos que el método que se aplicó para reducir la matriz aumentada del sistema (1.7) a su forma escalonada fue precisamente el de eliminación de Gauss. Véase (1.11).

Con base en lo anterior, podemos describir la idea de como resolver sistemas de ecuaciones lineales que venimos sugiriendo, así:

Idea básica para resolver un Sistema de Ecuaciones Lineales

1. Escriba la matriz aumentada del sistema de ecuaciones lineales.
2. Aplique el método de eliminación de Gauss para reducir la matriz aumentada del sistema a una matriz escalonada.
3. Resuelva el sistema correspondiente a la matriz escalonada por sustitución hacía atrás.

Teniendo en cuenta que, en un sistema de ecuaciones lineales cuya matriz asociada es una matriz escalonada, cada columna de la matriz tiene sólo un pivote o no tiene, llamamos *variables pivotales del sistema* a las variables del sistema correspondientes a las columnas pivotales de la matriz y *variables libres*[5] a las variables del sistema correspondientes a las columnas no pivotales de la matriz.

Ejemplo 12. Resolvamos el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & x_3 & = & 3 \\ 2x_1 & + & 3x_2 & + & x_3 & = & 5 \\ x_1 & - & x_2 & - & 2x_3 & = & -5 \end{array}. \tag{1.12}$$

1. Escribimos la matriz aumentada del sistema:

$$\left(\begin{array}{rrr|r} 2 & 1 & 1 & 3 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right).$$

2. Aplicamos el método de eliminación de Gauss a esta matriz:

$$\begin{array}{rcl} F_2 - F_1 & \longrightarrow & F_2 \\ F_3 - \frac{1}{2}F_1 & \longrightarrow & F_3 \end{array} \quad \left(\begin{array}{rrr|r} 2 & 1 & 1 & 3 \\ 0 & 2 & 0 & 2 \\ 0 & -3/2 & -5/2 & -13/2 \end{array}\right)$$

$$F_3 + \tfrac{3}{4}F_2 \longrightarrow F_3 \quad \left(\begin{array}{rrr|r} 2 & 1 & 1 & 3 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & -5/2 & -5 \end{array}\right).$$

Observemos que, en este ejemplo, x_1, x_2 y x_3 son las variables pivotales y que no hay variables libres.

3. Resolvemos el sistema correspondiente a la matriz aumentada escalonada,

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & x_3 & = & 3 \\ & & 2x_2 & & & = & 2 \\ & & & - & \frac{5}{2}x_3 & = & -5 \end{array} \qquad \begin{array}{c} \text{Aplicando} \\ \text{sustitución} \\ \text{hacia atrás} \end{array} \qquad \begin{array}{l} x_3 = 2 \\ x_2 = 1 \\ 2x_1 = 3 - x_2 - x_3 = 3 - 1 - 2 = 0. \\ x_1 = 0 \end{array}$$

Así, $(0, 1, 2)$ es la única solución del sistema (1.12). □

Ejemplo 13. Resolvamos el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcrcr} x & - & y & - & z & + & 2w & = & 1 \\ 2x & - & 2y & - & z & + & 3w & = & 3 \\ -x & + & y & - & z & & & = & -3 \end{array}. \tag{1.13}$$

1. Escribimos la matriz aumentada del sistema:

$$\left(\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array}\right).$$

[5]El nombre de variables libres está justificado en el hecho que, si el sistema es consistente, en su solución, esta variables pueden tomar valores arbitrarios. Además, observemos que este concepto de variable libre sólo tiene sentido en sistemas cuya matriz es escalonada.

2. Aplicamos el método de eliminación de Gauss a esta matriz:

$$\begin{array}{lcl} F_2 - 2F_1 & \longrightarrow & F_2 \\ F_3 + F_1 & \longrightarrow & F_3 \end{array} \quad \left(\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & -2 \end{array}\right)$$

$$F_3 + 2F_2 \longrightarrow F_3 \quad \left(\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Observemos que, en este ejemplo, x y z son las *variables pivotales* y que y y w son las *variables libres*.

3. Resolvemos el sistema correspondiente a la matriz aumentada escalonada,

$$\begin{array}{rcrcrcrcl} x & - & y & - & z & + & 2w & = & 1 \\ & & & & z & - & w & = & 1 \end{array} \quad \begin{array}{c} \text{Aplicando} \\ \text{sustitución} \\ \text{hacia atrás} \end{array} \quad \begin{array}{l} z = 1 + w \\ x = 1 + y + z - 2w \\ x = 1 + y + (1 + w) - 2w = 2 + y - w \end{array}$$

lo que nos indica que este sistema, y por tanto el sistema (1.13), tiene infinitas soluciones, ya que las variables y y w pueden tomar cualquier valor. □

Ejemplo 14. Ahora, resolvamos el sistema de ecuaciones lineales homogéneo asociado al sistema de ecuaciones anterior

$$\begin{array}{rcrcrcrcl} x & - & y & - & z & + & 2w & = & 0 \\ 2x & - & 2y & - & z & + & 3w & = & 0 \\ -x & + & y & - & z & & & = & 0 \end{array} . \qquad (1.14)$$

1. Escribimos la matriz aumentada del sistema:

$$\left(\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 0 \\ 2 & -2 & -1 & 3 & 0 \\ -1 & 1 & -1 & 0 & 0 \end{array}\right).$$

2. Aplicamos el método de eliminación de Gauss a esta matriz:

$$\begin{array}{lcl} F_2 - 2F_1 & \longrightarrow & F_2 \\ F_3 + F_1 & \longrightarrow & F_3 \end{array} \quad \left(\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & -2 & 2 & 0 \end{array}\right)$$

$$F_3 + 2F_2 \longrightarrow F_3 \quad \left(\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Observemos que, al igual que en el ejemplo anterior, x y z son *variables pivotales* y que y y w son *variables libres*.

3. Resolvemos el sistema correspondiente a la matriz aumentada escalonada,

$$\begin{array}{rcrcrcrcl} x & - & y & - & z & + & 2w & = & 0 \\ & & & & z & - & w & = & 0 \end{array} \quad \begin{array}{c} \text{Aplicando} \\ \text{sustitución} \\ \text{hacia atrás} \end{array} \quad \begin{array}{l} z = w \\ \\ x = y + z - 2w = y + w - 2w = y - w \end{array}$$

lo que nos indica que este sistema, y por tanto el sistema (1.14), tiene infinitas soluciones, $\{(x,\ y,\ z,\ w) : x = y - w,\ z = w,\ y, w \in R\}$. Recordemos que ésta no es la única forma de escribir este conjunto. □

Ejemplo 15. Resolvamos el sistema de ecuaciones

$$\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & 0 \\ x & + & 2y & - & z & = & 2 \\ & - & 5y & + & 5z & = & -1 \end{array}. \tag{1.15}$$

1. Escribimos la matriz aumentada del sistema:

$$\left(\begin{array}{rrr|r} 2 & -1 & 3 & 0 \\ 1 & 2 & -1 & 2 \\ 0 & -5 & 5 & -1 \end{array}\right).$$

2. Aplicamos el método de eliminación de Gauss a esta matriz:

$$F_2 - \tfrac{1}{2}F_1 \longrightarrow F_2 \quad \left(\begin{array}{rrr|r} 2 & -1 & 3 & 0 \\ 0 & 5/2 & -5/2 & 2 \\ 0 & -5 & 5 & -1 \end{array}\right)$$

$$F_3 + 2F_2 \longrightarrow F_3 \quad \left(\begin{array}{rrr|r} 2 & -1 & 3 & 0 \\ 0 & 5/2 & -5/2 & 2 \\ 0 & 0 & 0 & 3 \end{array}\right).$$

Observemos que, en este ejemplo, a diferencia de los anteriores, la columna de los términos independientes es una columna pivotal.

3. Al escribir el sistema correspondiente a la matriz aumentada escalonada,

$$\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & 0 \\ & & \frac{5}{2}y & - & \frac{5}{2}z & = & 2 \\ & & & & 0 & = & 3 \end{array},$$

obtenemos una ecuación que no tiene solución, 0=3, lo que nos indica que este sistema, y por tanto el sistema (1.15), no tiene solución. □

Analizando detenidamente los últimos 4 ejemplos, obtenemos las siguientes conclusiones:

1. Si al escalonar la matriz aumentada de un sistema de ecuaciones lineales, una fila tiene su pivote en la columna de los términos independientes, el sistema es inconsistente; o lo que es lo mismo, si al escalonar la matriz aumentada de un sistema de ecuaciones lineales, se tiene una fila de ceros en la matriz de los coeficientes y el término independiente de esta fila es diferente de cero, el sistema es inconsistente.

2. Si al escalonar la matriz aumentada de un sistema de ecuaciones lineales, no se obtiene pivote en la columna de los términos independientes, el sistema es consistente; o lo que es lo mismo, si al escalonar la matriz aumentada de un sistema de ecuaciones lineales, a todas las filas de ceros en la matriz de los coeficientes les corresponde un término independiente igual a cero, el sistema es consistente.

3. Si al escalonar la matriz aumentada de un sistema de ecuaciones lineales, se observan menos pivotes que columnas en la matriz de coeficientes y el sistema es consistente, el sistema tiene infinitas soluciones; o lo que es lo mismo, si el sistema es consistente y hay columnas sin pivote (variables libres), el sistema tiene infinitas soluciones.

4. Si al escalonar la matriz aumentada de un sistema de ecuaciones lineales, cada columna de la matriz de los coeficientes tiene pivote y el sistema es consistente, el sistema tiene solución única o lo que es lo mismo, si el sistema es consistente y no hay variables libres, el sistema tiene solución única.

5. Al escalonar la matriz aumentada de un sistema de ecuaciones lineales homogéneo, nunca habrá pivote en la columna de términos independientes ya que ésta siempre estará conformada por ceros. Por esta razón, para resolver un sistema de ecuaciones lineales homogéneo, basta escalonar la matriz de coeficientes del sistema de ecuaciones lineales en lugar de hacerlo con la matriz aumentada.

6. Si al escalonar la matriz de coeficientes de un sistema de ecuaciones lineales, cada fila tiene pivote, el sistema es consistente para cualquier columna de términos independientes.

De los ejemplos y las observaciones anteriores, es claro que para determinar si un sistema de ecuaciones lineales tiene o no solución y cuantas soluciones tiene, en caso de tenerlas, basta con llevar la matriz aumentada del sistema a una forma escalonada equivalente. En otras palabras, la consistencia o inconsistencia de un sistema de ecuaciones lineales y el número de soluciones se determina antes del Paso 3 de la idea básica para resolver un Sistema de Ecuaciones Lineales.

De otro lado, es fácil ver que despejar una variable en la sustitución hacia atrás es equivalente a convertir el pivote correspondiente en 1 mediante una operación elemental tipo escalamiento y que sustituir el valor obtenido de una variable en las ecuaciones anteriores es equivalente a introducir ceros arriba del pivote correspondiente mediante una operación elemental tipo eliminación, es decir que resolver un sistema con *patrón escalonado* mediante sustitución hacia atrás consiste en realizar operaciones elementales adecuadas a la matriz escalonada del sistema de ecuaciones lineales, para convertir en uno (1) los pivotes y obtener ceros encima de ellos, como se indica en el siguiente algoritmo.

Método de sustitución hacía atrás

Objetivo: Dado un sistema de ecuaciones lineales consistente con "patrón escalonado", encontrar su(s) solución(es).

1. Escriba la matriz aumentada del sistema de ecuaciones lineales.

2. Identifique la columna del primer pivote, de derecha a izquierda (ó la del último pivote, de izquierda a derecha).

3. Aplique la operación elemental adecuada de Tipo Escalamiento para obtener uno (1) como pivote en esta columna.

4. Aplique sucesivamente operaciones elementales adecuadas del Tipo Eliminación para obtener ceros encima del pivote de esta columna.

5. Repítale este procedimiento, comenzando con el Paso 2, al conjunto de filas de la matriz resultante en el paso anterior que están por encima de la fila donde estaba el pivote anterior, hasta agotar las filas.

6. Al final, si todas las columnas de la matriz de coeficientes tienen pivote, la solución es única y ésta aparecerá en las primeras filas de la columna de términos independientes correspondientes a las filas que tienen pivotes; en caso contrario (hay al menos una variable libre), el sistema tiene infinitas soluciones y se obtendrá despejando las variables pivotales en términos de las variables libres, lo cual será relativamente fácil.

Ejemplo 16. Usemos el método de sustitución hacia atrás, en forma matricial, en el Paso 3 del Ejemplo 12.

1. Después de aplicar el método de eliminación de Gauss, obtenemos la matriz escalonada (la cual es la matriz aumentada del sistema de ecuaciones lineales con "patrón escalonado").
$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & \mathbf{-5} & -10 \end{array}\right)$$

2. El primer pivote, de derecha a izquierda, es **-5**.

3. Obtengamos 1 como pivote, aplicando una operación elemental entre filas de Tipo Escalamiento

$$(-1/5)F_3 \longrightarrow F_3 \qquad \left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & \mathbf{1} & 2 \end{array}\right).$$

4. Obtengamos ceros encima del pivote, aplicando operaciones elementales entre filas de Tipo Eliminación

$$\begin{array}{ccc} F_2 + F_3 & \longrightarrow & F_2 \\ F_1 - F_3 & \longrightarrow & F_1 \end{array} \qquad \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right).$$

5. Repetimos los pasos con las filas por encima del pivote anterior. Esta vez, no necesitamos el Paso 3 ya que el pivote es **1**. Continuamos con el Paso 4 y obtenemos

$$F_1 - F_2 \longrightarrow F_1 \qquad \left(\begin{array}{ccc|c} 1 & 0 & 0 & \mathbf{0} \\ 0 & \mathbf{1} & 0 & \mathbf{1} \\ 0 & 0 & 1 & \mathbf{2} \end{array}\right).$$

6. Finalmente, todas las columnas de la matriz de coeficientes tienen pivote igual a 1, por lo tanto $(0, 1, 2)$ (componentes de la columna de términos independientes correspondientes a las filas con pivotes) es la única solución del sistema (1.12). □

Ejemplo 17. Usemos el método de sustitución hacia atrás, en forma matricial, en el Paso 2 del Ejemplo 13.

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & \mathbf{1} & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

1. El último pivote, de izquierda a derecha, es **1**. Pasamos directamente al Paso 4 del método de sustitución hacia atrás, y obtenemos

$$F_1 + F_2 \longrightarrow F_1 \qquad \left(\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

2. Finalmente, la segunda y cuarta columna no tienen pivote, por lo tanto y, w son variables libres y x, z son variables pivotales, las cuales procedemos a despejar en términos de las anteriores para obtener las infinitas soluciones del sistema (1.13).

$$\begin{array}{ccccccc} x - y & + w & = & 2 & \text{Despejando} & x = 2 + y - w \\ & & & & \text{las variables} & \\ z & - w & = & 1 & \text{pivotales} & z = 1 + w \end{array}$$

□

Finalmente, con este método para resolver sistemas de ecuaciones lineales con patrón escalonado, podemos presentar un algoritmo para la solución de un Sistema de Ecuaciones Lineales, que incluye el método de eliminación de Gauss y el de sustitución hacía atrás.

Método de solución de un Sistema de Ecuaciones Lineales(Eliminación de Gauss + Sustitución hacia atrás)
Objetivo: Dado un sistema de ecuaciones lineales, determinar si tiene solución y, en caso de tenerla, hallar su(s) solución(es)

1. Escriba la matriz aumentada del sistema de ecuaciones lineales.
2. Aplique el método de eliminación de Gauss para reducir la matriz aumentada del sistema a una matriz escalonada.
3. Si la columna de términos independientes es pivotal, el Sistema de Ecuaciones Lineales no tiene solución y termina el algoritmo. Si no, el Sistema de Ecuaciones Lineales tiene solución y continua el algoritmo.
4. Resuelva el sistema correspondiente a la matriz escalonada por el método de sustitución hacía atrás.

Una variante muy popularizada del método que acabamos de presentar es el llamado *Método de Gauss-Jordan* que consiste en convertir el pivote en 1, luego introducir ceros debajo del pivote y después introducir ceros encima de él, para cada columna pivotal. En otras palabras, este método sigue los pasos de los métodos de eliminación de Gauss y de sustitución hacía atrás **entremezclados**, el cual podemos resumir como sigue, teniendo en cuenta que, en general, este método no es mejor que el planteado anteriormente.

Otro método de solución de un Sistema de Ecuaciones Lineales(Gauss-Jordan)
Objetivo: Dado un sistema de ecuaciones lineales, determinar si tiene solución y, en caso de tenerla, hallar su(s) solución(es)

1. Escriba la matriz aumentada del sistema de ecuaciones lineales.
2. Identifique la primera columna, de izquierda a derecha, que no sea de sólo ceros.
3. Si el primer elemento de esta columna es un cero, intercambie la primera fila con una que tenga un elemento no cero en esta columna. Este elemento no cero será el pivote de esta columna.
4. Aplique la operación elemental adecuada de Tipo Escalamiento para obtener 1 como pivote.
5. Aplique sucesivamente operaciones elementales adecuadas de Tipo Eliminación, para obtener ceros debajo del pivote de esta columna.
6. Aplique sucesivamente operaciones elementales adecuadas de Tipo Eliminación para obtener ceros encima del pivote de esta columna.
7. Repítale los Pasos 2, 3 y 4, a las filas por debajo de la fila donde estaba el pivote anterior, y el Paso 5 a las filas por encima del nuevo pivote, hasta agotar las filas o hasta que las filas restantes sean de sólo ceros.
8. Si la columna de términos independientes es pivotal, el Sistema de Ecuaciones no tiene solución. Si no, el Sistema de Ecuaciones Lineales tiene solución. Si todas las columnas de la matriz de coeficientes tienen pivote, la solución es única y aparecerá en las primeras filas de la columna de términos independientes correspondientes a las filas que tienen pivotes; en caso contrario (hay al menos una variable libre), el sistema tiene infinitas soluciones y se obtendrá despejando las variables pivotales en términos de las variables libres, lo cual será relativamente fácil.

Aunque evidentemente, por este método también se llega a la solución del sistema, el alterar el orden de los

pasos del método de solución propuesto inicialmente (Eliminación de Gauss + Sustitución hacia atrás), hace que este otro método (Gauss-Jordan) sea ineficiente, puesto que aumenta el número de operaciones, casi al doble. Además, cuando el Sistema de Ecuaciones Lineales no tiene solución, se realizan muchas operaciones innecesarias antes de determinar esta situación. Es de anotar, que la computación (y/o programación) en paralelo elimina esta ineficiencia, ya que muchas de estas operaciones se pueden hacer simultáneamente produciendo un ahorro en el tiempo total de computo.

Ejemplo 18. Usando el Método de Gauss-Jordan, resolvamos el sistema de ecuaciones lineales del Ejemplo 11

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 3 \\ 2x_1 & + & 3x_2 & + & x_3 & = & 5 \\ x_1 & - & x_2 & - & 2x_3 & = & -5. \end{array} \tag{1.16}$$

1. Escribimos la matriz aumentada de este sistema:

$$\left(\begin{array}{rrr|r} \mathbf{1} & 1 & 1 & 3 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right).$$

2. La primer columna, de izquierda a derecha, diferente de cero es la Columna 1.

3. El primer pivote es **1**, por lo tanto no necesitamos los pasos 3 y 4 del Método de Gauss-Jordan. Obtenemos ceros debajo del pivote, aplicando operaciones elementales entre filas Tipo Eliminación

$$\begin{array}{rcl} F_2 - 2F_1 & \longrightarrow & F_2 \\ F_3 - F_1 & \longrightarrow & F_3 \end{array} \quad \left(\begin{array}{rrr|r} 1 & 1 & 1 & 3 \\ 0 & \mathbf{1} & -1 & -1 \\ 0 & -2 & -3 & -8 \end{array}\right)$$

4. Como es el primer pivote, no necesitamos el Paso 6 del Método de Gauss-Jordan. Repetimos los Pasos 2 y 3 del proceso con las filas debajo del anterior pivote (Filas 2 y 3), y encontramos que el pivote es nuevamente **1**, por lo tanto, no necesitamos los pasos 3 y 4 del Método de Gauss-Jordan. Obtenemos ceros debajo y encima del pivote, aplicando operaciones elementales entre filas Tipo Eliminación

$$\begin{array}{rcl} F_3 + 2F_2 & \longrightarrow & F_3 \\ F_1 - F_2 & \longrightarrow & F_1 \end{array} \quad \left(\begin{array}{rrr|r} 1 & 0 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & \mathbf{-5} & -10 \end{array}\right)$$

5. Repetimos los Pasos 2 y 3 del proceso con las filas debajo del pivote anterior (Fila 3) y encontramos que el pivote es **-5**, por lo tanto, aplicamos operación elemental entre filas Tipo Escalamiento para obtener **1** como pivote

$$(-1/5)F_3 \quad \longrightarrow \quad F_3 \quad \left(\begin{array}{rrr|r} 1 & 0 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & \mathbf{1} & 2 \end{array}\right)$$

6. Como es la última fila, solo necesitamos obtener ceros encima del pivote, aplicando operaciones elementales entre filas Tipo Eliminación

$$\begin{array}{rcl} F_2 + F_3 & \longrightarrow & F_2 \\ F_1 - 2F_3 & \longrightarrow & F_1 \end{array} \quad \left(\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & \mathbf{1} & 2 \end{array}\right)$$

7. Finalmente, todas las columnas de la matriz de coeficientes tienen pivote, por lo tanto, $(0, 1, 2)$ (filas de la columna de términos independientes correspondientes a las filas con pivotes) es la única solución del sistema. □

1.5. Solución Simultánea de Sistemas de Ecuaciones Lineales

Para terminar este capítulo, veamos que, cuando se requiere resolver dos ó más sistemas de ecuaciones lineales con la misma matriz de coeficientes y se tienen las columnas de términos independientes al mismo tiempo, es posible ahorrar tiempo de cálculo, resolviendo dichos sistemas simultáneamente.

En la sección anterior, vimos que las operaciones elementales usadas, tanto en el escalonamiento de la matriz aumentada de un sistema de ecuaciones lineales como en la sustitución hacia atrás, dependen exclusivamente, a menos que la matriz no tenga solución, de la matriz de coeficientes. Con base en este hecho, al resolver los sistemas de ecuaciones lineales cuyas matrices aumentadas son

$$\left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & b_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} & b_m \end{array}\right), \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & c_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} & c_m \end{array}\right), \ldots, \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & r_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & r_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} & r_m \end{array}\right),$$

podemos ahorrar tiempo de cálculo, aplicando el Método de Eliminación de Gauss + Sustitución hacia atrás, a la *matriz aumentada conjunta* de todos los sistemas,

$$\left(\begin{array}{cccc|cccc} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & b_1 & c_1 & \cdots & r_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & b_2 & c_2 & \cdots & r_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} & b_m & c_m & \cdots & r_m \end{array}\right)$$

en lugar de hacerlo a cada una de las matrices aumentadas correspondientes a cada sistema de ecuaciones lineales. Incluso, podemos ahorrar aún mas tiempo de computo, si antes de iniciar la sustitución hacia atrás, excluimos, de la matriz aumentada conjunta, las columnas correspondientes a los sistemas de ecuaciones lineales que no tengan solución. Ilustremos esta idea con el siguiente ejemplo.

Ejemplo 19: Resolvamos los sistemas de ecuaciones lineales

$$\begin{array}{rcrcrcrcr} & & & & z & + & 2w & = & -3 \\ 3x & - & 6y & - & 3z & & & = & 12 \\ -2x & + & 4y & + & & - & 4w & = & -2 \end{array},$$

$$\begin{array}{rcrcrcrcr} & & & & x_3 & + & 2x_4 & = & 0 \\ 3x_1 & - & 6x_2 & - & 3x_3 & & & = & -9 \\ -2x_1 & + & 4x_2 & + & & - & 4x_4 & = & 4 \end{array},$$

$$\begin{array}{rcrcrcrcr} & & & & z_3 & + & 2z_4 & = & -2 \\ 3z_1 & - & 6z_2 & - & 3z_3 & & & = & 3 \\ -2z_1 & + & 4z_2 & + & & - & 4z_4 & = & 2 \end{array}.$$

Como la matriz de coeficientes de los tres sistemas es la misma, entonces

1. Escribamos la matriz escalonada conjunta de estos tres sistemas de ecuaciones lineales

$$\left(\begin{array}{rrrr|rrr} 0 & 0 & 1 & 2 & -3 & 0 & -2 \\ 3 & -6 & -3 & 0 & 12 & -9 & 3 \\ -2 & 4 & 0 & -4 & -2 & 4 & 2 \end{array}\right).$$

2. A esta matriz, apliquémosle operaciones elementales entre filas para escalonarla (Eliminación de Gauss)

$$\begin{array}{ccc} F_1 & \longleftrightarrow & F_2 \\ F_3 + \frac{2}{3}F_1 & \longrightarrow & F_3 \end{array} \quad \left(\begin{array}{cccc|ccc} 3 & -6 & -3 & 0 & 12 & -9 & 3 \\ 0 & 0 & 1 & 2 & -3 & 0 & -2 \\ 0 & 0 & -2 & -4 & 6 & -2 & 4 \end{array}\right)$$

$$F_3 + 2F_2 \longrightarrow F_3 \quad \left(\begin{array}{cccc|ccc} 3 & -6 & -3 & 0 & 12 & -9 & 3 \\ 0 & 0 & 1 & 2 & -3 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & -2 & 0 \end{array}\right).$$

3. Como, en la matriz de coeficientes, la tercera fila es de ceros, determinemos cual(es) sistema(s) no tienen solución para excluirlos. Efectivamente, el segundo sistema no tiene solución, ya que en su correspondiente columna de términos independientes, su tercera componente es diferente de cero. Con esta misma matriz, también nos damos cuenta que la segunda y cuarta variables son variables libres, por lo tanto, los otros dos sistemas tienen infinitas soluciones. Excluyendo la segunda columna de términos independientes y la fila de ceros, tenemos.

$$\left(\begin{array}{cccc|cc} 3 & -6 & -3 & 0 & 12 & 3 \\ 0 & 0 & 1 & 2 & -3 & -2 \end{array}\right).$$

4. A esta matriz, apliquémosle operaciones elementales entre filas para hallar las soluciones del primer y tercer sistema (Sustitución hacia atrás).

$$F_1 + 3F_2 \longrightarrow F_1 \quad \left(\begin{array}{cccc|cc} 3 & -6 & 0 & 6 & 3 & -3 \\ 0 & 0 & 1 & 2 & -3 & -2 \end{array}\right)$$

$$\tfrac{1}{3}F_1 \longrightarrow F_1 \quad \left(\begin{array}{cccc|cc} 1 & -2 & 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & 2 & -3 & -2 \end{array}\right)$$

5. Finalmente, despejando las variables pivotales, obtenemos que el conjunto solución del primer sistema es

$$\{(1 + 2r - 2s,\ r,\ -3 - 2s,\ s) :\ r, s \in R\}$$

y el del tercer sistema es

$$\{(-1 + 2r - 2s,\ r,\ -2 - 2s,\ s) :\ r, s \in R\}.$$

□

Notemos que si los términos independientes de un sistema dependen de la solución de uno de los otros sistemas, ó si dichos términos no se tiene al mismo tiempo que los de los demás sistemas, el sistema correspondiente no se puede resolver simultáneamente con el resto. Para estas dos situaciones que muy frecuentemente aparecen en aplicaciones del mundo real, en la Sección 7 del Capitulo 3, presentaremos un esquema alternativo para ahorrar tiempo de computo.

1.6. Ejercicios

1. Dadas las siguientes ecuaciones:

(i) $4x_1 + 2x_2 + x_3 - 2 = 0$ (ii) $3x_1^2 - x_2 + 5x_3 = 2 - x_1$ (iii) $4x_1 - 3x_2 - 1 = x_3 + 5x_5$

(iv) $4x_1 + 2x_2 + x_3 = \dfrac{3x_2 + 5}{4 - 4x_3}$ (v) $\sqrt{3}x + \pi y - 12z = 7^{2/3}w$ (vi) $x_1 sen\frac{\pi}{3} + \frac{x_2}{4} - 3 = 0$

a) Cuales son *lineales*? Por qué?

b) Cuales son *homogéneas*? Por qué?

c) Cuales son los *coeficientes* y los *términos independientes* de las ecuaciones que usted identificó como lineales?

d) Cuales de las ternas $(-1, 2, 2)$, $(2, -1, 2)$, $(2, 2, -1)$ son *solución de la ecuación* (i)? Por qué?

e) Cuales de las 4-úplas $(1, 1, 1, 1)$, $(\sqrt{3}, 0, \frac{1}{4}, 0)$, $(12, 0, \sqrt{3}, 0)$ son *solución de la ecuación* (v)? Por qué?

f) Es posible encontrar otras soluciones de la ecuación (i)? De ser posible, halle al menos otra.

g) Es posible encontrar otras soluciones de la ecuación (v)? De ser posible, halle al menos otra.

h) Escriba las *ecuaciones homogéneas asociadas* a las ecuaciones que usted identificó como lineales.

2. Para cada una de las siguientes ecuaciones, determine sus variables con los respectivos coeficientes, su término independiente y su ecuación homogénea asociada. Determine además, si es posible, dos soluciones y el conjunto solución.

   (i) $4x + 2y - x - 3 = 0$ $\quad$ (ii) $3x - 2y + z = 2 - x + 2z$ $\quad$ (iii) $4x_1 - 3x_2 - 1 = x_3 + 5x_5$

3. Sean a, b números reales y x *la variable* de la ecuación $ax = b$.

   a) Es ésta una *ecuación lineal*?

   b) Es siempre posible hallar el valor de x (despejar x), que satisface la ecuación $ax = b$?

- Es posible encontrar una ecuación lineal que no tenga solución?
- Es posible encontrar una ecuación lineal homogénea que no tenga solución?

4. Encuentre todos los valores de a para los cuales cada una de las siguientes ecuaciones

   (i) $(a - 3)x = 5$ $\quad$ (ii) $2x = a$ $\quad$ (iii) $(a^2 - 4)x = 0$ $\quad$ (iv) $(a^2 - 4)x = a$

   a) tiene exactamente una solución.

   b) tiene infinitas soluciones.

   c) no tiene solución (es inconsistente).

5. Dados los siguientes sistemas de ecuaciones lineales:

   $(i)\ \begin{aligned} x - 2y &= 2 \\ 2x + 3y &= -3 \end{aligned}$ $\quad$ $(ii)\ \begin{aligned} x - 2y &= -1 \\ -2x + 4y &= 2 \end{aligned}$ $\quad$ $(iii)\ \begin{aligned} x - 2y &= -1 \\ -2x + 4y &= 0 \end{aligned}$ $\quad$ $(iv)\ \begin{aligned} x - 2y &= 0 \\ 3x - y &= 0 \end{aligned}$

   a) Cuales son *homogéneos*? Por qué?

   b) Cuales son los *coeficientes* y los *términos independientes* de los sistemas?

   c) Cuales de las duplas $(4, 1)$, $(0, -1)$, $(2, 2)$ son *solución del sistema* (i)? Por qué?

   d) Cuales de las duplas $(1, 1)$, $(-2, 4)$, $(3, 2)$ son *solución del sistema* (ii)? Por qué?

   e) Cuales de las duplas $(1, 1)$, $(-2, 4)$, $(3, 2)$ son *solución del sistema* (iii)? Por qué?

   f) Es posible encontrar otras soluciones del sistema (i)? De ser posible, halle al menos otra.

   g) Es posible encontrar otras soluciones del sistema (ii)? De ser posible, halle al menos otra.

   h) Determine cuales de estos sistemas son *consistentes*.

   i) Escriba los *sistemas de ecuaciones homogéneos asociados* a los sistemas de ecuaciones lineales dados.

6. Dado el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} 3x - 2y - z - 3 &= t \\ x + 3w - 2t &= 1 \\ -3x - y + 2z - 3w - t &= -2 \\ x + y + z - w &= 3 + 2t - w \end{aligned}$$

   a) Determine su matriz de coeficientes, sus términos independientes y su matriz aumentada. (AYUDA: defina un orden para las variables y reescriba cada una de las ecuaciones según este orden)

   b) Determine el sistema de ecuaciones lineales homogéneo asociado.

   c) Determine los pivotes, las variables pivotales y las variables libres.

7. Determine, en cada caso, si los sistemas de ecuaciones lineales dados son equivalentes:

   (i) $\begin{aligned} x - y &= 0 \\ -x - y &= 2 \end{aligned}$ y $\begin{aligned} 2u + 3v &= -5 \\ u - 2v &= 1 \end{aligned}$ (ii) $\begin{aligned} x - y &= 0 \\ -x - y &= 2 \end{aligned}$ y $\begin{aligned} x - y + z &= 0 \\ -x - y + z &= 2 \end{aligned}$

   (iii) $\begin{aligned} x_1 - x_2 &= 0 \\ -x_1 - x_2 &= 2 \\ -2x_1 + 2x_2 &= 0 \end{aligned}$ y $\begin{aligned} 2x_1 + 3x_2 &= -5 \\ x_1 - 2x_2 &= 1 \end{aligned}$

8. Dada la ecuación lineal (i) $4x + 2y + 3z = 2$, si multiplicamos ambos lados de la ecuación por 3, obtenemos la ecuación (ii) $12x + 6y + 9z = 6$.

   a) Pruebe que la terna $(1, \frac{1}{2}, -1)$ es una solución tanto de la ecuación (i), como de la ecuación (ii).

   b) Halle otra solución de la ecuación (ii) y muestre que también es solución de la ecuación (i).

   c) Demuestre que cualquier solución de la ecuación (i), es también solución de la ecuación (ii).

9. Dado el sistema de ecuaciones lineales

$$i) \quad \begin{aligned} x - y - z &= -1 \\ 3x - y + z &= 1 \end{aligned} ,$$

   si a la segunda ecuación del sistema le sumamos la primera multiplicada por -3, obtenemos el sistema de ecuaciones lineales

$$ii) \quad \begin{aligned} x - y - z &= -1 \\ 2y + 4z &= 4 \end{aligned} ,$$

   a) Verifique que la terna $(2, 4, -1)$ es una solución tanto del sistema de ecuaciones lineales (i), como del sistema de ecuaciones lineales (ii).

   b) Halle otra solución del sistema de ecuaciones lineales (i) y muestre que también es solución del sistema de ecuaciones lineales (ii).

   c) Demuestre que cualquier solución del sistema de ecuaciones lineales (ii) es también solución del sistema de ecuaciones lineales (i).

   d) Cuál de los tres resultados anteriores permiten concluir que los dos sistemas de ecuaciones son equivalentes?

   e) Encuentre las matrices aumentadas asociadas a los sistemas de ecuaciones lineales (i) y (ii). Verifique que estas matrices son equivalentes.

   f) Si dos sistemas de ecuaciones lineales son equivalentes, son sus matrices aumentadas asociadas equivalentes? (AYUDA: analice los sistemas iii) del ejercicio 7.)

10. Dada la matriz

$$A = \begin{pmatrix} 1 & -2 & 1 & 1 \\ 0 & 0 & -1/3 & 7/3 \\ 0 & 6 & -1 & -5 \end{pmatrix},$$

a) Escriba un sistema de ecuaciones lineales cuya matriz aumentada sea A.

b) Determine el sistema de ecuaciones lineales homogéneo asociado al sistema de ecuaciones lineales de la parte a).

11. Dados los siguientes sistemas de ecuaciones lineales, dibuje la gráfica correspondiente a cada uno de ellos. Determine geométricamente si cada sistema tiene solución y, en caso afirmativo, si la solución es única. Resuelva algebraícamente cada sistema para confirmar su respuesta.

(i) $\begin{aligned} x-2y &= 1 \\ 3x+y &= -4 \end{aligned}$ (ii) $\begin{aligned} x-2y &= -1 \\ -3x+6y &= 3 \end{aligned}$ (iii) $\begin{aligned} x-2y &= 1 \\ -3x+6y &= 0 \end{aligned}$

12. Dado el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcrcr} 2x & - & y & - & z & + & 2w & = & 0 \\ -4x & + & 2y & + & 3z & - & 3w & = & -1 \\ & - & 2y & - & z & + & 2w & = & 1 \end{array}$$

a) Efectue operaciones elementales entre las ecuaciones del sistema hasta llevarlo a uno con *patrón escalonado.*

b) Escriba la matriz de coeficientes y la matriz aumentada asociadas al sistema de ecuaciones lineales.

c) Describa las operaciones elementales entre filas correspondientes a las operaciones elementales entre ecuaciones efectuadas en la parte *a*).

13. Resuelva cada uno de los siguientes sistemas de ecuaciones lineales.

(i) $\begin{aligned} x+2y+3z &= 1 \\ -5y+2z &= 6 \\ 4z &= -8 \end{aligned}$ (ii) $\begin{aligned} x_1-3x_2+x_3 &= 5 \\ x_2-2x_3 &= -1 \end{aligned}$ (iii) $\begin{aligned} x &= -1 \\ -x+2y &= 5 \\ 3x+4y+2z &= 14 \end{aligned}$

14. Resuelva los sistemas de ecuaciones lineales encontrados en el ejercicio 10.

15. Los siguientes sistemas de ecuaciones son no lineales. Encuentre sustituciones de las variables que conviertan cada uno de ellos en un sistema de ecuaciones lineales y utilice este último para resolver el sistema inicialmente dado.

(i) $\begin{aligned} \frac{3}{x}+\frac{2}{y} &= 0 \\ \frac{4}{x}+\frac{3}{y} &= 1 \end{aligned}$ (ii) $\begin{aligned} -2^a+2(3^b) &= 1 \\ 3(2^a)-4(3^b) &= 1 \end{aligned}$ (iii) $\begin{aligned} x^2-y^2 &= 3 \\ x^2+2y^2 &= 6 \end{aligned}$

16. Demuestre que, para cada tipo de operación elemental entre filas, existe una operación elemental que **reversa** el procedimiento. En otras palabras, que si la matriz B resulta de aplicarle una operación elemental entre filas a la matriz A, existe otra operación elemental entre filas que al aplicársela a la matriz B se obtiene la matriz A.

17. Determine, para cada uno de los siguientes casos, si las matrices dadas son equivalentes (AYUDA: trate de encontrar una secuencia de operaciones elementales entre filas que al aplicárselas a una matriz se obtiene la otra)

(i) $\begin{pmatrix} -2 & 1 & 1 \\ 2 & 2 & -2 \end{pmatrix}$ *y* $\begin{pmatrix} -2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ (ii) $\begin{pmatrix} 1 & -1 \\ 2 & 1 \\ -3 & 2 \end{pmatrix}$ *y* $\begin{pmatrix} 1 & -1 \\ 0 & -1 \\ 0 & 3 \end{pmatrix}$

18. Determine cual o cuales de las siguientes matrices son escalonadas.

$(i)\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 3 & \sqrt{2} \\ 0 & 0 & \pi \end{pmatrix}$ $\quad (ii)\ \begin{pmatrix} 0 & 3 & 2 & 1 \\ 0 & 0 & 0 & -3 \end{pmatrix}$ $\quad (iii)\ \begin{pmatrix} 1 & 2 & 3 & 0 \end{pmatrix}$

19. A cada una de las siguientes matrices, aplique el método de eliminación de Gauss para encontrar una matriz escalonada equivalente a la matriz dada.

$(i)\ \begin{pmatrix} 0 & 2 \\ -1 & 3 \\ 7 & -1 \end{pmatrix}$ $\quad (ii)\ \begin{pmatrix} 2 & -1 & 4 \\ -4 & 2 & 0 \\ 0 & 3 & -1 \\ 2 & 1 & -3 \end{pmatrix}$ $\quad (iii)\ \begin{pmatrix} 0 & -1 & 3 & -1 \\ 5 & 2 & -1 & 2 \\ 10 & 3 & 0 & 3 \\ 5 & 1 & 1 & 1 \end{pmatrix}$

20. Dado el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcrcl} 2x & - & y & + & z & - & 2w & = & 0 \\ -4x & + & 2y & - & z & + & w & = & 3 \\ 2x & - & y & - & 2z & + & 3w & = & 2 \end{array},$$

a) Determine su matriz aumentada asociada.
b) Aplique el Algoritmo de Gauss para llevar la matriz encontrada en *a*) a una matriz escalonada.
c) Identifique los pivotes de la matriz escalonada.
d) Identifique las columnas pivotales de la matriz escalonada.
e) Identifique las variables pivotales del sistema.
f) Identifique las variables libres del sistema.
g) Determine el tipo (vacío, unitario o infinito) de conjunto solución del sistema.

21. Cada una de las siguientes matrices es una matriz escalonada equivalente a la matriz aumentada asociada a un sistema de ecuaciones lineales.

$(i)\ \begin{pmatrix} 2 & -1 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $\quad (ii)\ \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -3 \end{pmatrix}$ $\quad (iii)\ \begin{pmatrix} 2 & -1 & 4 & 1 & 0 \\ 0 & 1 & 2 & 3 & -2 \\ 0 & 0 & 0 & 2 & 0 \end{pmatrix}$

a) Determine el número de ecuaciones y el número de variables de cada uno de los sistemas.
b) Identifique las variables pivotales de cada uno de los sistemas.
c) Identifique las variables libres de cada uno de los sistemas.
d) Determine el tipo de conjunto solución de cada uno de los sistemas.
e) Para los sistemas consistentes, encuentre el conjunto solución.

22. Dados los siguientes sistemas de ecuaciones lineales:

$(i)\ \begin{array}{rcl} 3x - y & = & 1 \\ 2y + z & = & -2 \\ x - 2z & = & 3 \end{array}$ $\quad (ii)\ \begin{array}{rcl} 2x_1 - x_2 & = & x_4 \\ x_3 + x_4 & = & x_1 \\ x_4 - 2x_4 & = & x_2 \\ 2x_2 + x_3 & = & x_1 \end{array}$

a) Resuelva cada uno de ellos, aplicando el método de eliminación de Gauss y sustitución hacía atrás.
b) Resuelva cada uno de los sistemas de ecuaciones lineales homogéneos asociados.
c) Existe alguna relación entre los tipos de conjunto solución de un sistema de ecuaciones lineales y su sistema homogéneo asociado?

23. Para cada una de las siguientes situaciones, respecto del tamaño del sistema de ecuaciones lineales (número de ecuaciones × número de variables) y el número de variables pivotales, qué puede decirse sobre el tipo de conjunto solución del sistema?

Tamaño del sistema	No. de variables pivotales	Tamaño del sistema	No. de variables pivotales
3×4	3	5×5	3
3×4	4	5×5	4
3×4	2	5×5	5
4×3	3	8×5	5
4×3	4	8×5	4
4×3	2	5×8	6
5×5	2	5×8	4

Trate de generalizar sus respuestas para un sistema de ecuaciones lineales con m ecuaciones, n variables $(m \times n)$ y k variables pivotales.

24. Si al escalonar la matriz aumentada de un sistema de ecuaciones lineales, se obtiene

$$\left(\begin{array}{ccccc|c} \sqrt{2} & 0 & 3 & -1 & 4 & 0 \\ 0 & 2 & 2 & -\pi & -1 & 1 \\ 0 & 0 & 0 & a & 1 & 5 \\ 0 & 0 & 0 & 0 & b^2 - b & b \end{array}\right).$$

a) Es el sistema consistente cuando $a = b = 0$? En caso de serlo, es la solución única?

b) Es el sistema consistente cuando $a = 1$ y $b = 0$? En caso de serlo, es la solución única?

c) Es el sistema consistente cuando $a = 0$ y $b = 1$? En caso de serlo, es la solución única?

d) Si $b = 2$ y $a \neq 0$, qué puede decirse del conjunto solución?

e) Si $b = 1$ y $a \neq 0$, qué puede decirse del conjunto solución?

f) Si $a \neq 0$, dé un valor de b (diferente de 0), en caso de que exista, para que el sistema sea consistente.

g) Si $a \neq 0$, para que valores de b el sistema tiene infinitas soluciones?

h) Si $a \neq 0$, para que valores de b el sistema tiene solución única?

i) Si $a \neq 0$, para que valores de b el sistema es inconsistente?

25. Qué puede decirse del tipo de conjunto solución de un sistema de ecuaciones lineales donde una de sus ecuaciones es

a) $0 = 0$. b) $0 = 5$. c) $3 = 3$.

26. Conteste FALSO o VERDADERO y diga POR QUE, a cada una de las siguientes afirmaciones

a) Si un sistema de ecuaciones lineales tiene solución, cualquier otro sistema de ecuaciones lineales con la misma matriz de coeficiente también tiene solución.

b) Si un sistema de ecuaciones lineales tiene solución única, su sistema de ecuaciones lineales homogéneo asociado también tiene solución única.

c) Un sistema de ecuaciones lineales tiene solución siempre que su sistema de ecuaciones lineales homogéneo asociado tenga solución.

d) El tipo de conjunto solución de un sistema de ecuaciones lineales y el del sistema de ecuaciones lineales homogéneo asociado siempre es el mismo

e) Un sistema de ecuaciones lineales con 10 variables y 7 ecuaciones tiene infinitas soluciones.

f) Un sistema de ecuaciones lineales con 20 variables y 20 ecuaciones tiene solución única.

g) Un sistema de ecuaciones lineales con 14 variables y 10 ecuaciones no tiene solución única.

h) Un sistema de ecuaciones lineales con 27 variables y 13 ecuaciones puede no tener solución.

i) Un sistema de ecuaciones lineales con 100 variables y 300 ecuaciones puede tener solución única.

j) Un sistema de ecuaciones lineales homogéneo con 10 variables y 7 ecuaciones tiene infinitas soluciones.

k) Un sistema de ecuaciones lineales homogéneo con 20 variables y 20 ecuaciones tiene solución única.

l) Un sistema de ecuaciones lineales homogéneo con 14 variables y 10 ecuaciones no tiene solución única.

m) Un sistema de ecuaciones lineales homogéneo con 27 variables y 13 ecuaciones puede no tener solución.

n) Un sistema de ecuaciones lineales homogéneo con 100 variables y 300 ecuaciones puede tener solución única.

ñ) Un sistema de ecuaciones lineales consistente con 10 variables y 7 ecuaciones tiene infinitas soluciones.

o) Un sistema de ecuaciones lineales consistente con 20 variables y 20 ecuaciones tiene solución única.

p) Un sistema de ecuaciones lineales consistente con 14 variables y 10 ecuaciones no tiene solución única.

q) Un sistema de ecuaciones lineales consistente con 27 variables y 13 ecuaciones puede no tener solución.

r) Un sistema de ecuaciones lineales consistente con 100 variables y 300 ecuaciones puede tener solución única.

27. Resuelva, en caso de ser posible, los siguientes sistemas de ecuaciones lineales

a) $\begin{array}{rcl} 2x-4z & = & -1 \\ x+y-w & = & 1 \\ y+z-2w & = & 0 \end{array}$, $\begin{array}{rcl} 2x_1-4x_3 & = & 0 \\ x_1+x_2-x_4 & = & 2 \\ x_2+x_3-2x_4 & = & 4 \end{array}$, $\begin{array}{rcl} 2y_1-4y_3 & = & 6 \\ y_1+y_2-y_4 & = & 0 \\ y_2+y_3-2y_4 & = & -2 \end{array}$ y $\begin{array}{rcl} 2z_1-4z_3 & = & 3 \\ z_1+z_2-z_4 & = & -1 \\ z_2+z_3-2z_4 & = & 2 \end{array}$

b) $\begin{array}{rcl} x-2y+z & = & 1 \\ 2x-y-z & = & 0 \\ -3x+\ \ 3z & = & 1 \end{array}$, $\begin{array}{rcl} x_1-2x_2+x_3 & = & 0 \\ 2x_1-x_2-x_3 & = & 3 \\ -3x_1+\ \ 3x_3 & = & 2 \end{array}$ y $\begin{array}{rcl} u-2v+w & = & x_o \\ 2u-v-w & = & y_o \\ -3u+\ \ 3w & = & 0 \end{array}$

donde $(x_o, y_o, 0)$ es una solución del primer sistema. (AYUDA: Observe que los tres sistemas tienen la misma matriz de coeficientes).

Capítulo 2

VECTORES en R^n

2.1. Introducción

Una vez tenemos claro lo que es un Sistema de Ecuaciones Lineales y su representación matricial, el significado de su solución, el tipo de conjunto solución y un método para calcular y/o analizar dicho conjunto, queremos formalizar uno de los conceptos matemáticos que intuitivamente utilizamos en el capítulo anterior. En este capítulo, estudiaremos las n-úplas como vectores de R^n, las principales operaciones entre ellos y sus aplicaciones e interpretaciones geométricas, bien sean relacionados con un sistema de coordenadas o como vectores libres, particularizando a los casos de R^2 y R^3. Dado que una ecuación lineal con n variables representa geométricamente un hiperplano en R^n, siendo un punto, una recta y un plano casos particulares para n =1, 2 y a 3, respectivamente, aprovecharemos estas operaciones y sus propiedades algebraicas para hacer una introducción a la geometría analítica. También introducimos conceptos básicos del álgebra lineal tales como combinación lineal, independencia lineal, espacio generado y conjunto generador, lo cual esperamos sirva de base para su generalización en el capítulo de Espacios Vectoriales.

2.2. Conceptos Básicos

En el mundo real, existen cantidades como longitud, volumen, masa y temperatura que pueden determinarse sólo con su magnitud. Pero, también existen otras como la posición, la aceleración y la fuerza que además de la magnitud, requieren de una dirección y un sentido para determinarse; u otras como las notas de un semestre de un estudiante que cursa 6 materias y la producción de una empresa que tiene 10 líneas de productos, las cuales requieren de una lista ordenada de números. Estos últimos casos los conocemos como cantidades vectoriales a diferencia del primero que son cantidades escalares. Con base en esta diferencia intuitiva, veamos la definición de vector y sus operaciones básicas.

Definición 1 [*Vector de R^n*]. Llamamos *vector de R^n* a una lista ordenada de n números reales, la cual denotamos como $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. A x_1 lo llamamos *primera componente*, a x_2 *segunda componente* y en general, a x_k lo llamamos *k-ésima componente* del vector $\mathbf{x}$.

Cualquier vector cuyas componentes sean cero lo llamamos *vector nulo* o vector cero y lo denotamos $\mathbf{0}$.

Ejemplo 1. El vector $\begin{pmatrix} 5 \\ 0 \\ -3 \\ 5 \end{pmatrix}$ es un vector de R^4 y su primera, segunda, tercera y cuarta componentes

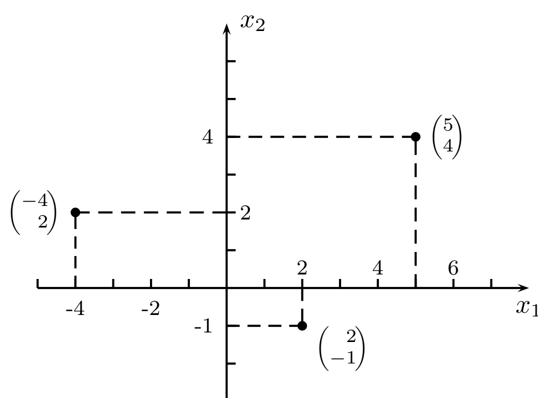
son 5, 0, -3 y 5, en ese orden. El vector $\begin{pmatrix} -1 \\ 1/5 \end{pmatrix}$ es un vector de R^2 y su primera y segunda componentes son -1 y 1/5, respectivamente. □

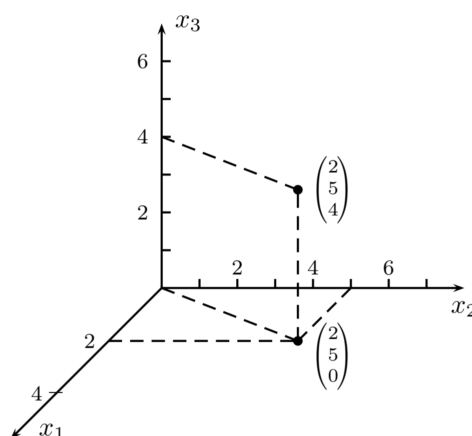
Ejemplo 2. Los vectores $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, $\cdots$, $\mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ son vectores de R^n. A estos vectores los llamamos *vectores canónicos* de R^n.[1] □

Diremos que dos vectores son *iguales* si todas sus componentes correspondientes son iguales. Así, para que el vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ sea igual al vector $\begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix}$, a debe ser -2, b debe ser 1 y c debe ser 3 y, por razones similares, el vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ es diferente del vector $\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$, ya que $1 \neq 3$.

Geométricamente, a los vectores de R^n los podemos interpretar como puntos; en particular, cuando $n = 2$ ó $n = 3$, son puntos del *plano* o del *espacio*, respectivamente.

Ejemplo 3. Representemos los vectores $\begin{pmatrix} -4 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 5 \\ 4 \end{pmatrix}$ *y* $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ en el plano (R^2) y los vectores $\begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix}$ *y* $\begin{pmatrix} 2 \\ 5 \\ 4 \end{pmatrix}$ en el espacio (R^3).





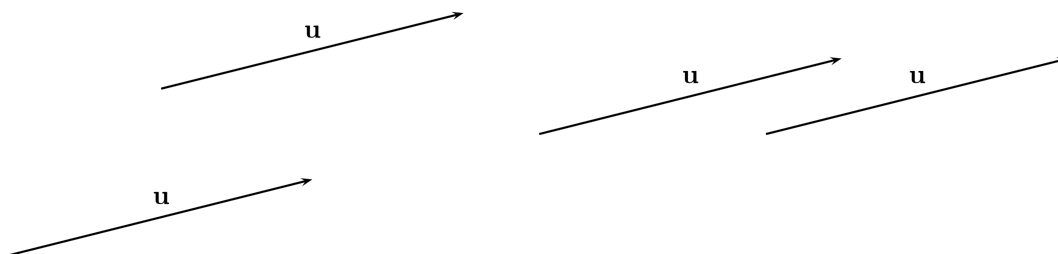
□

En las aplicaciones físicas, es importante que pensemos en un vector, no como un punto, sino como algo que tiene magnitud, dirección y sentido.

Definición 2 [*Vector libre*]. Un vector libre es un objeto matemático que podemos determinar por su dirección, sentido y magnitud.

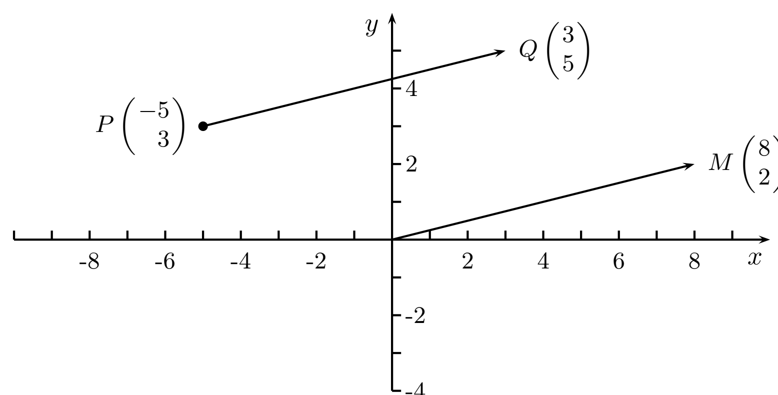
[1] Algunos autores denotan los vectores canónicos de R^3 como **i**, **j** y **k**, respectivamente; es decir, $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ y $\mathbf{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Geométricamente, podemos representar los vectores libres como *segmentos dirigidos*[2], los que denotamos con letras en negrilla, tales como **u**, **v**, etc., o bien como $\overline{PQ}$, $\overline{RS}$, donde las primeras letras corresponden a los puntos iniciales y las otras a los puntos finales de estos segmentos. Si dos segmentos dirigidos $\overline{PQ}$ y $\overline{RS}$ tienen igual magnitud, dirección y sentido, decimos que son *iguales*. De esta forma, para cada vector dado, hay infinitos vectores libres iguales a él, como lo muestra la gráfica siguiente.



Al vector cuyo punto inicial coincide con su punto final lo llamamos *vector nulo o vector* **0**. Este vector no tiene dirección, ni sentido y su magnitud es 0.

Queremos resaltar que, dado un sistema de coordenadas (o sistema de referencia), podemos representar geométricamente un vector como un punto y un vector libre como un segmento dirigido. Así, dado un vector libre $\overline{PQ}$, existe un vector $\overline{OM} = \overline{PQ}$, donde O es el origen del sistema de coordenadas y M es el punto cuyas coordenadas son la diferencia de las coordenadas de los puntos Q y P. Observemos que las coordenadas del vector $\overline{OM}$ coinciden con las del punto M. La siguiente sección nos dará más elementos para entender esta observación.



2.3. Operaciones con Vectores

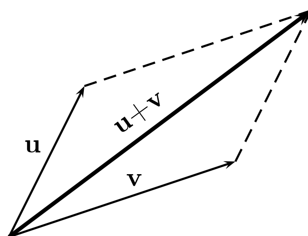
Al igual que con las cantidades escalares, dadas unas cantidades vectoriales, es deseable obtener otras a partir de ellas. Para empezar, esto es posible usando las operaciones básicas entre vectores: la suma y la multiplicación por escalar que definiremos a continuación.

Definición 3 [*Suma de vectores*]. Definimos la *suma entre dos vectores* **u** y **v** de R^n como el vector **u**+**v**, cuyas

[2]Llamamos segmento dirigido $\overline{PQ}$ a la parte de una recta comprendida entre dos puntos: uno inicial (P) y uno final (Q), su longitud representa la magnitud del vector, la recta que lo contiene define la dirección del vector y la orientación de P hacía Q representa el sentido.

componentes son la suma de las componentes respectivas de los vectores **u** y **v**; es decir, dados $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$, definimos $\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$.

En término de vectores libres, el vector $\mathbf{u} + \mathbf{v}$ es el vector que va desde el punto inicial del vector **u** hasta el punto final del vector **v**, después de *mover paralelamente* el vector **v** de tal manera que su punto inicial coincida con el punto final de **u**. En otras palabras, el vector $\mathbf{u} + \mathbf{v}$ es la diagonal del paralelogramo cuyos lados son **u** y **v**, como lo ilustramos en la siguiente gráfica.

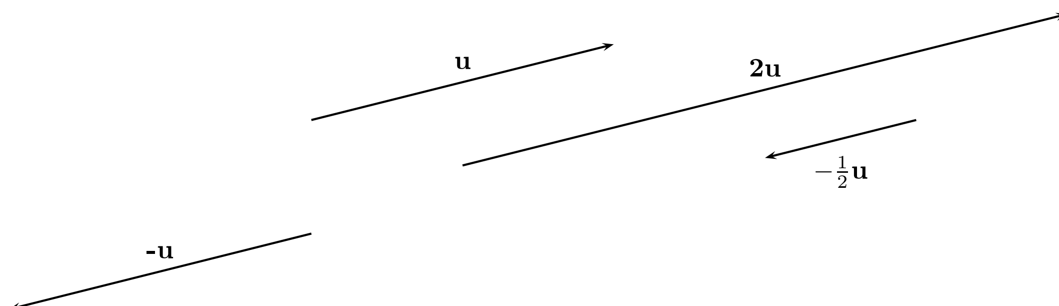


Definición 4 [*Producto por escalar*]. Definimos el *producto por escalar* de un vector **u** por un número real λ como el vector $\lambda\mathbf{u}$, cuyas componentes son el producto de λ por las componentes respectivas del vector **u**; es decir, dados $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ y $\lambda \in R$, definimos $\lambda\mathbf{u} = \begin{pmatrix} \lambda u_1 \\ \lambda u_2 \\ \vdots \\ \lambda u_n \end{pmatrix}$.

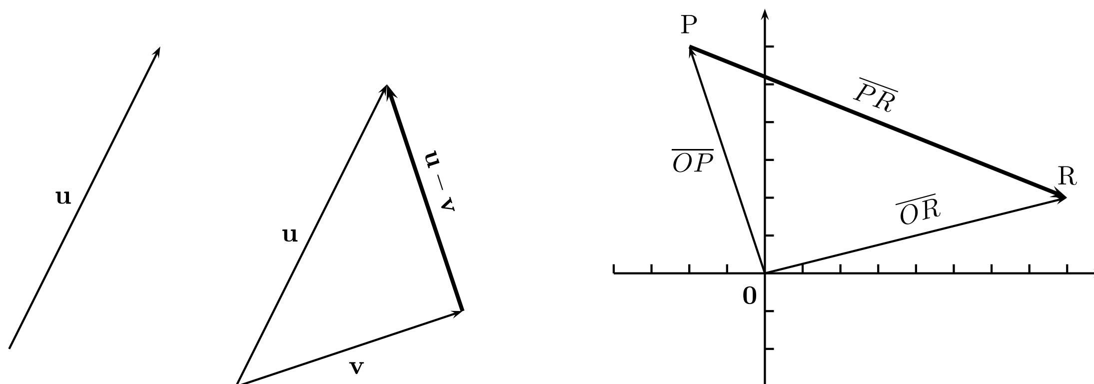
En término de vectores libres, el vector $\lambda\mathbf{u}$ es el vector que tiene igual dirección que el vector **u** y que dependiendo del signo de λ, tiene igual sentido ($\lambda>0$) o sentido opuesto ($\lambda<0$) al del vector **u**, y cuya magnitud es $|\lambda|$ por la magnitud del vector **u**.

Un caso especial de esta operación es $(-1)\mathbf{u}$, lo cual denotamos como $-\mathbf{u}$ y lo denominamos *vector opuesto* a **u**, ya que tiene la misma dirección y magnitud que **u** pero sentido opuesto.

De la interpretación geométrica de la multiplicación por escalar, deducimos que dos vectores distintos de cero *tienen igual dirección (o son paralelos)*, si y solo si, el uno es un escalar por el otro. Así, los vectores de la siguiente figura son paralelos ya que son múltiplos por escalar unos de otros.



Definida la suma y el producto por escalar, podemos definir la *resta* $\mathbf{u} - \mathbf{v}$ como la suma de $\mathbf{u}$ con el opuesto de $\mathbf{v}$. En término de vectores libres, la resta $\mathbf{u} - \mathbf{v}$ es la diagonal del paralelogramo de lados $\mathbf{u}$ y $\mathbf{v}$ que va desde el punto final de $\mathbf{v}$ hasta el punto final de $\mathbf{u}$, cuando ellos tienen el mismo punto inicial. Así, al vector $\overline{PR}$ lo podemos ver como la resta $\overline{OR} - \overline{OP}$, tal como lo ilustramos en la gráfica y lo habíamos mencionado al final de la sección anterior.



Muchos de los cálculos que hacemos con vectores son similares a los que efectuamos con escalares ó números reales, pero ojo! No siempre es así, como veremos más adelante. Para llamar la atención sobre este punto, veamos cuáles son las propiedades algebraicas que poseen las dos operaciones básicas antes definidas, las cuales son muy parecidas a las que poseen los números reales.

Teorema 1 [*Propiedades de la suma y el producto por escalar de vectores*].
Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores de R^n y sean α y β dos números reales. Entonces se cumplen las siguientes proposiciones:

1. $(\mathbf{u} + \mathbf{v}) \in R^n$. *Ley clausurativa para la suma*
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$. *Ley asociativa para la suma*
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. *Ley conmutativa para la suma*
4. Existe un único vector $\mathbf{z} \in R^n$ tal que $\mathbf{u} + \mathbf{z} = \mathbf{z} + \mathbf{u} = \mathbf{u}$ $(\mathbf{z} = \mathbf{0})$. *Ley modulativa para la suma*
5. Para cada $\mathbf{u}$, existe un único vector $\mathbf{p} \in R^n$ tal que $\mathbf{u} + \mathbf{p} = \mathbf{p} + \mathbf{u} = \mathbf{0}$ $(\mathbf{p} = -\mathbf{u})$. *Existencia del opuesto para la suma*
6. $\lambda\mathbf{u} \in R^n$. *Ley clausurativa para el producto por escalar*
7. $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$. *Ley distributiva del producto por escalar respecto a la suma de vectores*
8. $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$. *Ley distributiva del producto por escalar respecto a la suma de escalares*
9. $(\alpha\beta)\mathbf{u} = \alpha(\beta\mathbf{u}) = \beta(\alpha\mathbf{u})$. *Ley asociativa respecto al producto por escalares*
10. $1\mathbf{u} = \mathbf{u}$. *Ley modulativa para el producto por escalar*
11. $0\mathbf{u} = \mathbf{0}$.
12. $\alpha\mathbf{0} = \mathbf{0}$.
13. $\alpha\mathbf{u} = \mathbf{0}$, si y solo si, $\alpha = 0$ ó $\mathbf{u} = \mathbf{0}$.

Demostración: La Propiedad 1 es inmediata de la definición de la suma de vectores. Demostremos la Propiedad 2 y las demás las dejamos como ejercicio para el lector. Sean u_i, v_i y w_i las i-ésimas componentes de los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$, respectivamente. Entonces, $u_i + v_i$ es la i-ésima componente de $\mathbf{u} + \mathbf{v}$, así que $(u_i + v_i) + w_i$ es la i-ésima componente de $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$. Del mismo modo, tenemos que $u_i + (v_i + w_i)$ es la i-ésima componente de $\mathbf{u} + (\mathbf{v} + \mathbf{w})$. Como $(u_i + v_i) + w_i = u_i + (v_i + w_i)$ para todo $i = 1, 2, \cdots, n$, por la propiedad asociativa de la suma de números reales, las componentes respectivas de $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$ y de $\mathbf{u} + (\mathbf{v} + \mathbf{w})$ son iguales. Concluimos entonces que $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$ y $\mathbf{u} + (\mathbf{v} + \mathbf{w})$ son iguales. □

El siguiente ejemplo ilustra la utilidad del teorema anterior.

Ejemplo 4.
$$\begin{aligned}(2\mathbf{u} - 3\mathbf{v} + \mathbf{w}) - (5\mathbf{u} - 2\mathbf{v}) + 7\mathbf{u} &= 2\mathbf{u} - 3\mathbf{v} + \mathbf{w} - 5\mathbf{u} + 2\mathbf{v} + 7\mathbf{u} \\ &= (2\mathbf{u} - 5\mathbf{u} + 7\mathbf{u}) + (-3\mathbf{v} + 2\mathbf{v}) + \mathbf{w} \\ &= (2 - 5 + 7)\mathbf{u} + (-3 + 2)\mathbf{v} + \mathbf{w} \\ &= 4\mathbf{u} - \mathbf{v} + \mathbf{w}\ .\end{aligned}$$

□

2.4. Combinación Lineal y Conjuntos Generado y Generador

Con las operaciones básicas, a partir de un conjunto de vectores, podemos obtener muchos vectores más: todos los múltiplos por escalar de los vectores iniciales y todas las sumas de estos. A continuación, presentamos la definición formal de estos elementos.

Definición 5 [*Combinación lineal*]. Dados $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vectores de R^n y $\lambda_1, \lambda_2, \ldots, \lambda_k \in R$, al vector

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k$$

lo llamamos *combinación lineal* de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. A los escalares $\lambda_1, \lambda_2, \ldots, \lambda_k$ los llamamos *coeficientes* de la combinación lineal.

Ejemplo 5. Sean $\mathbf{u} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$ y $\mathbf{w} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$. Calculemos la combinación lineal de ellos dada por $3\mathbf{u} - \mathbf{v} + 2\mathbf{w}$.

$$\begin{aligned}3\mathbf{u} - \mathbf{v} + 2\mathbf{w} &= 3\begin{pmatrix} -1 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ 5 \end{pmatrix} + 2\begin{pmatrix} 3 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} -3 \\ 6 \end{pmatrix} + \begin{pmatrix} -2 \\ -5 \end{pmatrix} + \begin{pmatrix} 6 \\ -4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ -3 \end{pmatrix}.\end{aligned}$$

□

Ejemplo 6. Determinemos si los vectores $\begin{pmatrix} -13 \\ 4 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$ son combinaciones lineales de $\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$ y $\begin{pmatrix} -5 \\ 2 \\ -3 \end{pmatrix}$; es decir, veamos si existen escalares α, β, λ y μ, tales que

$$\alpha\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + \beta\begin{pmatrix} -5 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} -13 \\ 4 \\ 0 \end{pmatrix} \quad \text{y} \quad \lambda\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + \mu\begin{pmatrix} -5 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

Equivalentemente, veamos si los sistemas

$$\begin{array}{rcl} \alpha - 5\beta &=& -13 \\ 2\beta &=& 4 \\ -2\alpha - 3\beta &=& 0 \end{array} \qquad \text{y} \qquad \begin{array}{rcl} \lambda - 5\mu &=& 2 \\ 2\mu &=& 0 \\ -2\lambda - 3\mu &=& -1 \end{array}$$

son consistentes. Aplicando las operaciones elementales $F_3 + 2F_1 \to F_3$ y $F_3 + \frac{13}{2}F_2 \to F_3$ a la matriz aumentada conjunta de los sistemas anteriores,

$$\left(\begin{array}{rr|rr} 1 & -5 & -13 & 2 \\ 0 & 2 & 4 & 0 \\ -2 & -3 & 0 & -1 \end{array}\right),$$

obtenemos la matriz escalonada conjunta

$$\left(\begin{array}{rr|rr} 1 & -5 & -13 & 2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 3 \end{array}\right),$$

lo que indica que el primer sistema es consistente, mientras que el segundo no lo es. Así que $\begin{pmatrix} -13 \\ 4 \\ 0 \end{pmatrix}$ es combinación lineal de $\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$ y $\begin{pmatrix} -5 \\ 2 \\ -3 \end{pmatrix}$, mientras que $\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$ no lo es.

□

Del ejemplo anterior, observemos que para determinar si un vector es combinación lineal de otros vectores no es necesario calcular los escalares de la combinación lineal; basta con verificar que existen.

El conjunto de todas las combinaciones lineales de un conjunto de vectores dado, por su importancia práctica y teórica, recibe un nombre especial.

Definición 6 [*Conjunto generado y Conjunto generador*]. Al conjunto de todas las combinaciones lineales de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ lo llamamos *conjunto generado* por los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y lo representamos por

$$Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \{\mathbf{v} \mid \mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k, \quad \lambda_i \in R\}.$$

En otras palabras, si $V = Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, decimos que V es generado por $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$; además, a $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ lo llamamos *conjunto generador* de V.

Ejemplo 7. Si $V = Gen\{\mathbf{u}, \mathbf{v}\}$, entonces $2\mathbf{u} + \sqrt{5}\mathbf{v}, \mathbf{0}, \mathbf{u}, 3\mathbf{v}, \mathbf{u} - \mathbf{v}$ son vectores de V.

Es claro que $2\mathbf{u} + \sqrt{5}\mathbf{v}$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$, lo que implica que $2\mathbf{u} + \sqrt{5}\mathbf{v} \in V$. Veamos que cada uno de los otros vectores también es combinación lineal de $\mathbf{u}$ y $\mathbf{v}$.

$$\begin{array}{rcl} \mathbf{0} &=& 0\mathbf{u} + 0\mathbf{v} \\ \mathbf{u} &=& 1\mathbf{u} + 0\mathbf{v} \end{array} \qquad \begin{array}{rcl} 3\mathbf{v} &=& 0\mathbf{u} + 3\mathbf{v} \\ \mathbf{u} - \mathbf{v} &=& 1\mathbf{u} + (-1)\mathbf{v}. \end{array}$$

Por lo tanto pertenecen a V. □

Ejemplo 8. El conjunto generado por un solo vector es el conjunto de todos los múltiplos por escalar de él; es decir, es el conjunto de todos los vectores paralelos al vector dado.

En efecto, si $V = Gen\{\mathbf{v}\}$ y $\mathbf{u} \in V$, entonces $\mathbf{u} = \lambda\mathbf{v}$ para algún $\lambda \in R$ y si $\mathbf{w} = \alpha\mathbf{v}$ para algún $\alpha \in R$, entonces $\mathbf{w} \in V$. □

Ejemplo 9. Demostremos que $Gen\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} = R^n$, donde $\mathbf{e}_i$ es el i-ésimo vector canónico de R^n.

Es claro que cualquier combinación lineal de los vectores $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ es un vector de R^n. Veamos ahora que cualquier vector de R^n lo podemos escribir como combinación lineal de los vectores $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$.

En efecto, sea $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \in R^n$, entonces $\mathbf{u} = u_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + u_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$.[3] □

Ejemplo 10. El vector $\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}$ pertenece a $Gen \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right\}$.

Veamos que efectivamente existen escalares α y β tales que $\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$. O lo que es equivalente, que existen escalares α y β tales que

$$\begin{aligned} 3 &= 1\alpha + 0\beta \\ 0 &= 0\alpha + 0\beta \\ 2 &= 1\alpha - 1\beta, \end{aligned}$$

para lo cual basta escalonar la matriz $\left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 0 & 0 \\ 1 & -1 & 2 \end{array}\right)$.

Así, realizando las operaciones elementales $F_3 - F_1 \to F_3$ y $F_2 \leftrightarrow F_3$, obtenemos $\left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{array}\right)$; de donde, podemos concluir que efectivamente el sistema tiene solución y por tanto $\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} \in Gen \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right\}$.

Adicionalmente, si queremos encontrar la combinación lineal explícita, tenemos que resolver el sistema asociado a la matriz escalonada,

$$\begin{aligned} 1\alpha + 0\beta &= 3 \\ 0\alpha - 1\beta &= -1, \end{aligned}$$

obteniendo $\alpha = 3$ y $\beta = 1$. Así que $\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$. □

2.5. Producto Ax

Observemos que las columnas de las matrices introducidas en el primer capítulo son vectores de R^m, siendo m el número de filas de la matriz; es decir, que las matrices las podemos ver como una sucesión (finita y ordenada) de vectores, lo cual denotamos como

$$A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \ldots \ \mathbf{a}_n]$$

[3] En particular, para $n = 3$, si $\mathbf{u} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, entonces $\mathbf{u} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$

para indicar que los vectores $\mathbf{a}_1,\ \mathbf{a}_2,\ \ldots,\ \mathbf{a}_n$, en su orden, son las columnas de la matriz A. Esta notación nos permite simplificar la escritura de las combinaciones lineales estudiadas en la sección anterior, como lo planteamos en la siguiente definición.

Definición 7 [*Producto* $A\mathbf{x}$]. Sea A una matriz, cuyas columnas son los vectores $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ de R^m, y sea $\mathbf{x}$ un vector de R^n, cuyas componentes son $x_1, x_2, \ldots, x_n$. Definimos el producto matricial $A\mathbf{x}$ como la combinación lineal

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n.$$

Observemos que si multiplicamos la matriz A por el vector canónico $\mathbf{e}_j$, el resultado es la j-ésima columna de A; en otras palabras,

$$A\mathbf{e}_j = \mathbf{a}_j\ .$$

Ejemplo 11. Dados $A = \begin{pmatrix} -1 & 0 & 3 \\ 2 & 1 & 1 \\ 3 & 5 & -2 \end{pmatrix}$ y $\mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$, encontremos $A\mathbf{x}$.

$$A\mathbf{x} = 0\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix} + 3\begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 9 \\ 4 \\ -1 \end{pmatrix}.$$

□

El siguiente teorema describe las propiedades fundamentales de esta operación.

Teorema 2 [*Propiedades del producto* $A\mathbf{x}$].
Dados A, una matriz cuyas columnas son los vectores $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ de R^m, λ un escalar y $\mathbf{x}$, $\mathbf{y}$ vectores de R^n, entonces

1. $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$
2. $A(\lambda\mathbf{x}) = \lambda(A\mathbf{x})$

Demostración: Demostremos la primera propiedad, dejando la segunda como ejercicio para el lector. Sean $x_1, x_2, \ldots, x_n$ y $y_1, y_2, \ldots, y_n$ las componentes de los vectores $\mathbf{x}$ y $\mathbf{y}$, respectivamente. Entonces, dado que las componentes del vector $(\mathbf{x}+\mathbf{y})$ son $x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n$, y aplicando las propiedades algebraicas de la suma de vectores y del producto por escalar (Teorema 1), tenemos

$$\begin{aligned} A(\mathbf{x} + \mathbf{y}) &= (x_1 + y_1)\mathbf{a}_1 + (x_2 + y_2)\mathbf{a}_2 + \ldots + (x_n + y_n)\mathbf{a}_n \\ &= x_1\mathbf{a}_1 + y_1\mathbf{a}_1 + x_2\mathbf{a}_2 + y_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n + y_n\mathbf{a}_n \\ &= (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n) + (y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + \ldots + y_n\mathbf{a}_n) \\ &= A\mathbf{x} + A\mathbf{y} \end{aligned}$$

□

Notemos que el producto $A\mathbf{x}$ nos ofrece una forma elegante y simple de representar un sistema de ecuaciones lineales.

Ejemplo 12. El sistema

$$\begin{aligned} 3x - 2y + z &= -2 \\ x - 3z &= 1 \end{aligned}$$

lo podemos expresar en forma vectorial como

$$x\begin{pmatrix} 3 \\ 1 \end{pmatrix} + y\begin{pmatrix} -2 \\ 0 \end{pmatrix} + z\begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

y esta ecuación vectorial, teniendo en cuenta la definición del producto de una matriz por un vector, lo podemos escribir como

$$\begin{pmatrix} 3 & -2 & 1 \\ 1 & 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Dicho de otro modo, un sistema de ecuaciones lineales lo podemos expresar en la forma $A\mathbf{x} = \mathbf{b}$, donde A representa la matriz de los coeficientes, $\mathbf{x}$ el vector de las incógnitas y $\mathbf{b}$ el vector de los términos independientes. De esta manera, tenemos que $[A|\mathbf{b}]$, $A\mathbf{x} = \mathbf{b}$ y $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n = \mathbf{b}$ son distintas formas de representar un sistema de ecuaciones lineales.

Así, afirmar que, por ejemplo, un sistema $[A|\mathbf{b}]$ es consistente, es equivalente a decir que hay un vector $\mathbf{x}$, tal que $\mathbf{b}$ lo podemos expresar como el producto $A\mathbf{x}$. De aquí que las siguientes afirmaciones son equivalentes, cuyas demostraciones dejamos para el lector.

Teorema 3 [*Equivalencia de conceptos*].
Dados A una matriz, cuyas n columnas son vectores de R^m y $\mathbf{b}$ un vector de R^m, las siguientes proposiciones son equivalentes:

1. El sistema cuya matriz aumentada es $[A|\mathbf{b}]$ es consistente.
2. Existe al menos un vector $\mathbf{x}$ de R^n, tal que $A\mathbf{x} = \mathbf{b}$.
3. El vector $\mathbf{b}$ es combinación lineal de las columnas de A.
4. El vector $\mathbf{b}$ pertenece al conjunto generado por las columnas de A.

□

Al conjunto de vectores $\mathbf{x}$ de R^n, tales que $A\mathbf{x} = \mathbf{0}$ y al conjunto de vectores $\mathbf{b}$, tales que $A\mathbf{x} = \mathbf{b}$ para algún vector $\mathbf{x}$ de R^n, los llamaremos de manera especial, por las propiedades que tienen.

Definición 8 [*Espacio nulo*]. Dada A, una matriz con n columnas, definimos *el espacio nulo* de A como el conjunto N_A de todos los vectores $\mathbf{x}$ de R^n, tales que $A\mathbf{x} = \mathbf{0}$. Esto es

$$N_A = \{\mathbf{x} \in R^n : A\mathbf{x} = \mathbf{0}\}.$$

Ejemplo 13. Dada $A = \begin{pmatrix} -1 & 2 & 1 \\ 2 & 1 & -1 \end{pmatrix}$, determinemos si los vectores $\mathbf{u} = \begin{pmatrix} -2 \\ 7 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$ y $\mathbf{w} = \begin{pmatrix} -3 \\ 1 \\ -5 \end{pmatrix}$ se encuentran en N_A.

Rápidamente nos damos cuenta que $\mathbf{u}$ no está en N_A, ya que $A\mathbf{u}$ no está definido, y que $\mathbf{v}$ tampoco, porque

$$A\mathbf{v} = \begin{pmatrix} 0 \\ 7 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Mientras que $\mathbf{w}$ si se encuentra en N_A, ya que

$$A\mathbf{w} = -3\begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 5\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

□

Observemos que $\mathbf{0}$ siempre está en N_A (el vector $\mathbf{0}$ siempre es solución de los sistemas de ecuaciones lineales homogéneos) y que $N_A = \{\mathbf{0}\}$, si y solo si, $A\mathbf{x} = \mathbf{0}$ tiene solución única. Las propiedades más importantes del espacio nulo están contenidas en el siguiente teorema.

Teorema 4 [*Propiedades del espacio nulo*].
Dada una matriz A, si $\mathbf{x}$, $\mathbf{y}$ son vectores de N_A y λ es un escalar, tenemos que:

1. $\mathbf{x} + \mathbf{y} \in N_A$
2. $\lambda\mathbf{x} \in N_A$

Demostración: Puesto que $\mathbf{x}, \mathbf{y} \in N_A$, $A\mathbf{x} = \mathbf{0}$ y $A\mathbf{y} = \mathbf{0}$; entonces,

1. $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, por tanto, $\mathbf{x} + \mathbf{y} \in N_A$.
2. $A(\lambda\mathbf{x}) = \lambda(A\mathbf{x}) = \lambda\mathbf{0} = \mathbf{0}$, de donde, concluimos que $\lambda\mathbf{x} \in N_A$.

□

Definición 9 [*Espacio columna*]. Dada A, una matriz con n vectores columna de R^m, definimos *el espacio columna* de A como el conjunto C_A de todos los vectores $\mathbf{b}$ de R^m para los que existe un vector $\mathbf{x}$ de R^n tal que $A\mathbf{x} = \mathbf{b}$. Esto es

$$C_A = \{\mathbf{b} \in R^m : A\mathbf{x} = \mathbf{b}, \text{ para algún } \mathbf{x} \in R^n\}.$$

Observemos que C_A está formado por todas las combinaciones lineales de las columnas de A; es decir,

$$C_A = Gen\{\mathbf{a}_1, \ \mathbf{a}_2, \ \ldots, \ \mathbf{a}_n\},$$

donde $\mathbf{a}_1, \ \mathbf{a}_2, \ \ldots, \ \mathbf{a}_n$ son las columnas de A.

Ejemplo 14. Dada $A = \begin{pmatrix} -1 & 2 \\ 2 & 1 \\ 1 & -1 \end{pmatrix}$, determinemos si $\begin{pmatrix} 4 \\ 7 \\ -1 \end{pmatrix}$ y $\begin{pmatrix} 5 \\ 0 \\ -4 \end{pmatrix}$ se encuentran en C_A.

Este problema es equivalente, por el Teorema 3, a determinar si los sistemas cuya matriz aumentada conjunta es

$$\left(\begin{array}{rr|rr} -1 & 2 & 4 & 5 \\ 2 & 1 & 7 & 0 \\ 1 & -1 & -1 & -4 \end{array}\right)$$

son consistentes. Al escalonar esta matriz, obtenemos

$$\left(\begin{array}{rr|rr} -1 & 2 & 4 & 5 \\ 0 & 5 & 15 & 10 \\ 0 & 0 & 0 & -1 \end{array}\right),$$

lo que nos indica que el primer sistema es consistente, mientras que el segundo no lo es. Así que $\begin{pmatrix} 4 \\ 7 \\ -1 \end{pmatrix}$ se encuentra en C_A, mientras que $\begin{pmatrix} 5 \\ 0 \\ -4 \end{pmatrix}$ no.

□

Las dos propiedades fundamentales del espacio columna están contenidas en el siguiente teorema.

Teorema 5 [*Propiedades del espacio columna*].
Dada una matriz A, los vectores $\mathbf{b}$ y $\mathbf{c}$ de C_A y λ un escalar, entonces:

1. $\mathbf{b} + \mathbf{c} \in C_A$
2. $\lambda\mathbf{b} \in C_A$

Demostración: Puesto que $\mathbf{b}, \mathbf{c} \in C_A$, existen vectores $\mathbf{x}$ y $\mathbf{y}$ tales que $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{y} = \mathbf{c}$. Entonces,

1. $\mathbf{b} + \mathbf{c} = A\mathbf{x} + A\mathbf{y} = A(\mathbf{x} + \mathbf{y})$, por tanto, $\mathbf{b} + \mathbf{c} \in C_A$.
2. $\lambda\mathbf{b} = \lambda(A\mathbf{x}) = A(\lambda\mathbf{x})$, de donde concluimos que $\lambda\mathbf{b} \in C_A$. □

A raíz de las propiedades anteriores, observemos las relaciones que existen entre las soluciones de un sistema y las de su sistema homogéneo asociado.

***Corolario 5.1*.**
Dada una matriz A, si el vector $\mathbf{u}$ es solución del sistema $A\mathbf{x} = \mathbf{b}$ y el vector $\mathbf{v}$ es solución del sistema homogéneo asociado ($A\mathbf{x} = \mathbf{0}$), entonces $(\mathbf{u} + \mathbf{v})$ es solución del sistema $A\mathbf{x} = \mathbf{b}$.

Demostración:

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

□

***Corolario 5.2*.**
Dada una matriz A, si los vectores $\mathbf{u}$ y $\mathbf{v}$ son soluciones del sistema $A\mathbf{x} = \mathbf{b}$, entonces $\mathbf{u} - \mathbf{v}$ es solución del sistema homogéneo asociado $A\mathbf{x} = \mathbf{0}$.

Demostración: Si los vectores $\mathbf{u}$ y $\mathbf{v}$ son soluciones del sistema $A\mathbf{x} = \mathbf{b}$, entonces $A\mathbf{u} = \mathbf{b}$ y $A\mathbf{v} = \mathbf{b}$. Restando estas dos últimas igualdades, tenemos que $A(\mathbf{u} - \mathbf{v}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$. Así que $(\mathbf{u} - \mathbf{v})$ es solución de $A\mathbf{x} = \mathbf{0}$. □

Estos dos últimos resultados nos permiten caracterizar las soluciones de un sistema de ecuaciones lineales, conociendo una solución de él y el conjunto solución del sistema homogéneo asociado.

***Corolario 5.3*.**
Dada una matriz A y una solución $\mathbf{u}$ del sistema $A\mathbf{x} = \mathbf{b}$, $\mathbf{v}$ es solución del sistema $A\mathbf{x} = \mathbf{b}$, si y solo si, $\mathbf{v} = \mathbf{h} + \mathbf{u}$, donde $\mathbf{h}$ es una solución del sistema homogéneo asociado $A\mathbf{x} = 0$.

Demostración: Sea $\mathbf{v}$ una solución del sistema $A\mathbf{x} = \mathbf{b}$, entonces $\mathbf{h} = \mathbf{v} - \mathbf{u}$ es solución del sistema homogéneo asociado (Corolario 5.2) y por tanto $\mathbf{v} = \mathbf{h} + \mathbf{u}$. La otra implicación es el resultado del Corolario 5.1. □

Con base en los tres últimos corolarios, podemos establecer que el número de soluciones de un sistema de ecuaciones lineales es cero, uno o infinito, lo cual resumimos en el siguiente corolario.

***Corolario 5.4*.**
Un sistema $A\mathbf{x} = \mathbf{b}$ que tiene más de una solución, tiene infinitas soluciones.

Demostración: Sean $\mathbf{u}$ y $\mathbf{v}$ dos soluciones diferentes del sistema $A\mathbf{x} = \mathbf{b}$. Por el Corolario 5.2, $\mathbf{h} = \mathbf{u} - \mathbf{v} \neq \mathbf{0}$ es solución del sistema homogéneo asociado $A\mathbf{x} = 0$. Por el Teorema 4, $\alpha\mathbf{h}$, para todo $\alpha \in R$, también es solución del sistema homogéneo, lo que nos indica que el sistema homogéneo tiene infinitas soluciones. Por el Corolario 5.3, $\mathbf{w} = \alpha\mathbf{h} + \mathbf{u}$ es también solución del sistema $A\mathbf{x} = \mathbf{b}$. Así que, el sistema $A\mathbf{x} = \mathbf{b}$ tiene infinitas soluciones. □

2.6. Independencia Lineal

El hecho que $\mathbf{w}$ pueda escribirse como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, podríamos interpretarlo como que $\mathbf{w}$ "depende" de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y, en este contexto, podemos decir que si un vector se puede expresar como combinación lineal de otros, éste depende de ellos. Pero, por ejemplo, si

$$\mathbf{w} = \frac{1}{2}\mathbf{u} - 3\mathbf{v},$$

también tenemos que

$$\mathbf{u} = 2\mathbf{w} + 6\mathbf{v},$$

lo que da origen a la pregunta de si es $\mathbf{w}$ el que depende de $\mathbf{u}$ y $\mathbf{v}$ o es $\mathbf{u}$ el que depende de $\mathbf{w}$ y $\mathbf{v}$. Para evitar este aparente inconveniente, establecemos la siguiente definición.

Definición 10 [*Conjunto de vectores linealmente dependientes*]. Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es *linealmente dependiente (l.d.)* si existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_k$, al menos uno de ellos diferente de cero, tales que

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k = \mathbf{0}.$$

Un conjunto de vectores que no es linealmente dependiente lo llamamos *linealmente independiente (l.i.)*. Es decir, un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es *linealmente independiente* si los únicos escalares $\lambda_1, \lambda_2, \ldots, \lambda_k$ tales que

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k = \mathbf{0}$$

son todos cero.

Ejemplo 15. Demostremos que $\left\{ \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ -5 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \right\}$ es un conjunto de vectores *l.i.*

Si $\lambda_1 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ -5 \\ 4 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, tenemos que el sistema correspondiente, además de ser homogéneo, tiene como matriz de coeficientes a $A = \begin{pmatrix} 1 & -1 & 1 \\ 3 & -5 & -2 \\ -2 & 4 & 0 \end{pmatrix}$, y que, al escalonar esta matriz, obtenemos $\begin{pmatrix} 1 & -1 & 1 \\ 0 & -2 & -5 \\ 0 & 0 & -3 \end{pmatrix}$. De esta última matriz, podemos concluir que el sistema de ecuaciones tiene como solución única el vector cero y por tanto, el conjunto de vectores es *l.i.*

□

Ejemplo 16. Demostremos que $\left\{ \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -5 \end{pmatrix} \right\}$ es un conjunto de vectores *l.d.*

Si $\lambda_1 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} + \lambda_3 \begin{pmatrix} 2 \\ 1 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, tenemos que el sistema correspondiente, además de ser homogéneo, tiene como matriz de coeficientes a $A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ -2 & 3 & -5 \end{pmatrix}$, y que, al escalonar esta matriz, obtenemos $\begin{pmatrix} 1 & -1 & 1 \\ 0 & 5 & -5 \\ 0 & 0 & 0 \end{pmatrix}$, de donde podemos concluir que el sistema de ecuaciones tiene infinitas soluciones y por tanto, el conjunto de vectores es *l.d.*

□

Ejemplo 17. Demostremos que todo conjunto que contenga el vector nulo es un conjunto de vectores *l.d.*

Es claro que

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k + \lambda_{k+1}\mathbf{0} = \mathbf{0},$$

cuando $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$ y λ_{k+1} toma cualquier valor; por tanto, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{0}\}$ es *l.d.*

□

Como indicaron los ejemplos anteriores, la independencia lineal de vectores la podemos determinar analizando

la forma escalonada de la matriz cuyas columnas son los vectores en cuestión, como lo establece el siguiente teorema.

Teorema 6 [*Equivalencia de conceptos*].
Dados los vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de R^m, sea A la matriz cuyas columnas son $\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_n$. Las siguientes afirmaciones son equivalentes.

1. El conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es *l.i.*
2. El sistema homogéneo $A\mathbf{x} = \mathbf{0}$ tiene solución única (la solución trivial o vector cero).
3. La forma escalonada de la matriz A tiene n pivotes.

Demostración: Para demostrar la equivalencia de estas tres afirmaciones, es suficiente demostrar que cada una de ellas implica la siguiente y que la última implica la primera. A continuación, demostramos que la primera afirmación implica la segunda y dejamos la demostración de las otras implicaciones como ejercicio para el lector.

Supongamos que el conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es *l.i.*, entonces la única combinación lineal de ellos igual a cero es la trivial. Usando la definición del producto $A\mathbf{x}$, lo anterior significa que el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ tiene solución única: el vector cero. □

2.7. Producto Escalar

Volviendo atrás, decíamos que un vector es un objeto matemático con magnitud, dirección y sentido; aspectos que intuitivamente entendemos, pero que deberíamos formalizar más. Para este efecto, introduciremos otra operación básica, el producto escalar entre dos vectores, la cual está relacionada con el concepto de magnitud de un vector y la noción de ángulo entre vectores.

Definición 11 [*Producto escalar*][4]. Dados dos vectores $\mathbf{u}$ y $\mathbf{v}$ de R^n, definimos $\mathbf{u}\cdot\mathbf{v}$, el *producto escalar* entre $\mathbf{u}$ y $\mathbf{v}$, como el escalar que obtenemos al sumar los productos de las componentes de $\mathbf{u}$ con las respectivas componentes de $\mathbf{v}$. En otras palabras, dados $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$, definimos

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n.$$

Ejemplo 18. Dados los vectores $\mathbf{u} = \begin{pmatrix} -2 \\ 2 \\ 0 \\ -5 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 1 \\ 12 \\ -3 \\ 0 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 3 \\ 1 \\ 5 \\ 4 \end{pmatrix}$ y $\mathbf{z} = \begin{pmatrix} -1 \\ 22 \\ 21 \end{pmatrix}$, calculemos los productos escalares entre ellos.

$$\mathbf{u}\cdot\mathbf{v} = -2\cdot 1 + 2\cdot 12 + 0\cdot(-3) + (-5)\cdot 0 = -2 + 24 + 0 + 0 = 22$$
$$\mathbf{u}\cdot\mathbf{w} = -2\cdot 3 + 2\cdot 1 + 0\cdot 5 + (-5)\cdot 4 = -6 + 2 + 0 - 20 = -24$$
$$\mathbf{v}\cdot\mathbf{u} = 1\cdot(-2) + 12\cdot 2 + (-3)\cdot 0 + 0\cdot 0 = -2 + 24 + 0 + 0 = 22$$
$$\mathbf{v}\cdot\mathbf{w} = 1\cdot 3 + 12\cdot 1 + (-3)\cdot 5 + 0\cdot 4 = 3 + 12 - 15 + 0 = 0$$
$$\mathbf{w}\cdot\mathbf{u} = 3\cdot(-2) + 1\cdot 2 + 5\cdot 0 + 4\cdot(-5) = -6 + 2 + 0 - 20 = -24$$
$$\mathbf{w}\cdot\mathbf{v} = 3\cdot 1 + 1\cdot 12 + 5\cdot(-3) + 4\cdot 0 = 3 + 12 - 15 + 0 = 0$$

[4]En la literatura, otros nombres usados para esta operación son *producto punto* y *producto interno*

Observemos que $\mathbf{u}\cdot\mathbf{v}=\mathbf{v}\cdot\mathbf{u}$, $\mathbf{u}\cdot\mathbf{w}=\mathbf{w}\cdot\mathbf{u}$ y $\mathbf{v}\cdot\mathbf{w}=\mathbf{w}\cdot\mathbf{v}$ y los productos escalares $\mathbf{u}\cdot\mathbf{z}$, $\mathbf{v}\cdot\mathbf{z}$ y $\mathbf{w}\cdot\mathbf{z}$ no están definidos. Por qué?

□

La pregunta que inmediatamente surge es ¿Cuáles de las propiedades del producto entre números reales se satisfacen para el producto escalar entre vectores? y su respuesta nos la da el siguiente teorema, cuya demostración dejamos como ejercicio para el lector.

Teorema 7 [*Propiedades del producto escalar*].
Dados los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ de R^n y el escalar α, tenemos que

1. $\mathbf{u}\cdot\mathbf{v}=\mathbf{v}\cdot\mathbf{u}$. *Ley conmutativa*
2. $\mathbf{u}\cdot(\mathbf{v}+\mathbf{w})=\mathbf{u}\cdot\mathbf{v}+\mathbf{u}\cdot\mathbf{w}$. *Ley distributiva para la suma de vectores*
3. $(\alpha\mathbf{u})\cdot\mathbf{v}=\alpha(\mathbf{u}\cdot\mathbf{v})$.

□

Ejemplo 19. Dados los vectores $\mathbf{u}=\begin{pmatrix}2\\1\\-5\end{pmatrix}$, $\mathbf{v}=\begin{pmatrix}1\\3\\0\end{pmatrix}$ y $\mathbf{w}=\begin{pmatrix}-2\\-1\\-1\end{pmatrix}$, calculemos $\mathbf{u}\cdot\mathbf{v}$, $\mathbf{u}\cdot\mathbf{w}$, $\mathbf{v}\cdot\mathbf{w}$, $(3\mathbf{u})\cdot\mathbf{v}$, $(\mathbf{u}+\mathbf{v})\cdot\mathbf{w}$ y $\mathbf{v}\cdot\mathbf{u}$.

Usando la definición y aplicando las propiedades del producto escalar, tenemos

$$\mathbf{u}\cdot\mathbf{v}=2\cdot1+1\cdot3+(-5)\cdot0=5$$
$$\mathbf{u}\cdot\mathbf{w}=2\cdot(-2)+1\cdot(-1)+(-5)\cdot(-1)=0$$
$$\mathbf{v}\cdot\mathbf{w}=1\cdot(-2)+3\cdot(-1)+0\cdot(-1)=-5$$
$$(3\mathbf{u})\cdot\mathbf{v}=3(\mathbf{u}\cdot\mathbf{v})=3\times5=15$$
$$(\mathbf{u}+\mathbf{v})\cdot\mathbf{w}=\mathbf{u}\cdot\mathbf{w}+\mathbf{v}\cdot\mathbf{w}=0+(-5)=-5$$
$$\mathbf{v}\cdot\mathbf{u}=\mathbf{u}\cdot\mathbf{v}=5$$

□

Observemos que, a diferencia de lo que ocurre con el producto entre números reales, $\mathbf{u}\cdot\mathbf{v}=0$ no implica que $\mathbf{u}=\mathbf{0}$ o $\mathbf{v}=\mathbf{0}$, como lo muestra el segundo producto del ejemplo anterior. Las siguientes son otras propiedades del producto escalar, cuyas implicaciones geométricas veremos más adelante.

Teorema 8 [*Propiedades del producto escalar*].
Si $\mathbf{u}$ es un vector de R^n, entonces:

1. $\mathbf{u}\cdot\mathbf{u}\geq0$.
2. $\mathbf{u}\cdot\mathbf{u}=0$, si y solo si, $\mathbf{u}=\mathbf{0}$.

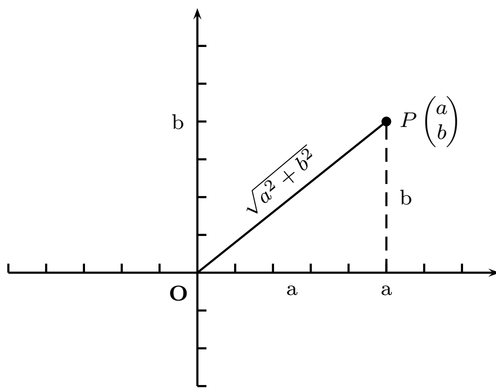
Demostración:

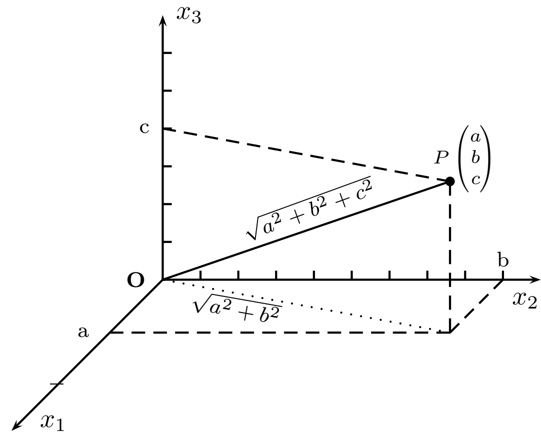
1. $\mathbf{u}\cdot\mathbf{u}=u_1u_1+u_2u_2+\ldots+u_nu_n=u_1^2+u_2^2+\ldots+u_n^2$ y la suma de cuadrados de números reales siempre es mayor o igual a cero.
2. $\mathbf{u}\cdot\mathbf{u}=u_1^2+u_2^2+\ldots+u_n^2=0$, si y solo si, $u_i^2=0$, para todo $i=1,\ldots n$. De donde concluimos que $\mathbf{u}\cdot\mathbf{u}=0$, si y solo si, $\mathbf{u}=\mathbf{0}$.

□

Como introducción a una de las aplicaciones geométricas del producto interno, notemos que si O es el origen de un sistema de coordenadas y P es un punto cuyas coordenadas son $\mathbf{p}=\begin{pmatrix}a\\b\end{pmatrix}$, aplicando el Teorema de

Pitágoras, la magnitud del segmento dirigido $\overline{OP}$, hipotenusa del triángulo rectángulo cuyos otros dos lados tienen magnitudes a y b, es $\sqrt{a^2+b^2} = \sqrt{\mathbf{p}\cdot\mathbf{p}}$, como vemos en la figura siguiente (izquierda). De la misma forma, si estamos en el espacio y P es un punto cuyas coordenadas son $\mathbf{p} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, aplicando el Teorema de Pitágoras dos veces, obtenemos que la magnitud del segmento dirigido $\overline{OP}$ es $\sqrt{a^2+b^2+c^2} = \sqrt{\mathbf{p}\cdot\mathbf{p}}$, como vemos en la figura siguiente (derecha).





De otro lado, tenemos que el segmento $\overline{OP}$ lo podemos ver como un vector, asi que la magnitud al cuadrado del vector $\overline{OP}$ en R^2 o R^3 es la suma de los cuadrados de las componentes del vector. Esta observación, nos induce a una definición de magnitud de un vector.

Definición 12 [*Norma*][5]. Definimos $\|\mathbf{u}\|$, la *norma del vector* $\mathbf{u}$ de R^n, como la raíz cuadrada de $\mathbf{u}\cdot\mathbf{u}$; es decir,

$$\|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}} = \sqrt{u_1^2 + \ldots + u_n^2}.$$

Ejemplo 20. Dados el vector $\mathbf{u} = \begin{pmatrix} 2 \\ 1 \\ -5 \end{pmatrix}$ y los puntos $P = \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix}$ y $Q = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$, calculemos $\|\mathbf{u}\|$ y $\|\overline{PQ}\|$.

$\|\mathbf{u}\| = \sqrt{2^2+1^2+(-5)^2} = \sqrt{30} \approx 5,477$ y $\|\overline{PQ}\| = \|Q-P\| = \sqrt{(1-5)^2+(-1-2)^2+(3-3)^2} = \sqrt{25} = 5.$ □

Las propiedades algebraicas de la norma de un vector, que deducimos de manera inmediata de los Teoremas 7 y 8, están contenidas en el siguiente teorema.

Teorema 9 [*Propiedades de la norma*].
Dados los vectores $\mathbf{u}$ y $\mathbf{v}$ y el escalar α, tenemos que

1. $\|\alpha\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$
2. $\|\mathbf{u}\| = 0$, si y solo si, $\mathbf{u} = \mathbf{0}$

Demostración:

1. $\|\alpha\mathbf{u}\| = \sqrt{\alpha\mathbf{u}\cdot\alpha\mathbf{u}} = \sqrt{\alpha^2(\mathbf{u}\cdot\mathbf{u})} = \sqrt{\alpha^2}\sqrt{\mathbf{u}\cdot\mathbf{u})} = |\alpha|\|\mathbf{u}\|$

[5] En la literatura, aparecen otros nombres, como *magnitud* y *longitud*, para referirse a la norma de un vector.

2. $\|\mathbf{u}\| = 0$, si y solo si, $\mathbf{u}\cdot\mathbf{u} = 0$, si y solo si, $\mathbf{u} = \mathbf{0}$. □

Decimos que el vector **u** es *unitario* cuando $\|\mathbf{u}\| = 1$. Es fácil ver que, si **u** es un vector no nulo, $\frac{1}{\|\mathbf{u}\|}\mathbf{u}$ es un vector unitario, que por ser múltiplo escalar positivo de **u**, tiene la misma dirección y sentido que **u**. ($-\frac{1}{\|\mathbf{u}\|}\mathbf{u}$ es otro vector unitario en la dirección de **u**. Existe otro?)

Ejemplo 21. Encontremos un vector unitario en la dirección de $\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$.

Dado que $\|\mathbf{v}\| = \sqrt{1^2 + (-1)^2 + 3^2} = \sqrt{11}$, tenemos que $\mathbf{u} = \frac{1}{\sqrt{11}}\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{11} \\ -1/\sqrt{11} \\ 3/\sqrt{11} \end{pmatrix}$ es un vector unitario en la dirección de **v**, ya que, por ser un múltiplo escalar de **v**, tiene su misma dirección y su norma es 1,

$$\left\| \begin{pmatrix} 1/\sqrt{11} \\ -1/\sqrt{11} \\ 3/\sqrt{11} \end{pmatrix} \right\| = \sqrt{(1/\sqrt{11})^2 + (-1/\sqrt{11})^2 + (3/\sqrt{11})^2} = \sqrt{1/11 + 1/11 + 9/11} = 1.$$

□

Ejemplo 22. Los vectores canónicos de R^n, presentados en el Ejemplo 2, tienen norma 1. □

Una pregunta que seguramente nos surge es ¿La norma de una suma o diferencia de vectores es la suma o diferencia de las normas de los vectores? y la respuesta es afirmativa sólo en un caso particular; en general, es sólo válida una desigualdad entre estas dos cantidades (*desigualdad triangular*). Para demostrar este resultado, son necesarios los dos siguientes teoremas.

Teorema 10 [*Norma de la suma y la resta de vectores*].
Dados los vectores **u** y **v** de R^n, tenemos que

1. $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2(\mathbf{u}\cdot\mathbf{v})$
2. $\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2(\mathbf{u}\cdot\mathbf{v})$

Demostración: Aplicando las propiedades del producto escalar (Teorema 7), tenemos

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \\ &= \mathbf{u}\cdot(\mathbf{u}+\mathbf{v}) + \mathbf{v}\cdot(\mathbf{u}+\mathbf{v}) \\ &= \mathbf{u}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{v}\ . \\ &= \mathbf{u}\cdot\mathbf{u} + 2(\mathbf{u}\cdot\mathbf{v}) + \mathbf{v}\cdot\mathbf{v} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) \end{aligned}$$

De manera similar, podemos demostrar la segunda propiedad (tomando $-\mathbf{v}$ en lugar de **v**). □

Teorema 11 [*Desigualdad de Cauchy-Schwarz*].
Dados los vectores **u** y **v** de R^n, tenemos que

$$|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\ \|\mathbf{v}\|.$$

Tenemos la igualdad, si y solo si, $\mathbf{u} = \lambda\mathbf{v}$ para algún $\lambda \in R$; es decir, si y sólo si, **u** y **v** son paralelos.

Demostración: Por las propiedades del producto escalar (Teorema 7), tenemos que, para todo $x \in R$,

$$\begin{aligned} 0 &\leq (x\mathbf{u}+\mathbf{v})\cdot(x\mathbf{u}+\mathbf{v}) \\ &= x^2\ (\mathbf{u}\cdot\mathbf{u}) + x\ (2\mathbf{u}\cdot\mathbf{v}) + (\mathbf{v}\cdot\mathbf{v}) = p(x), \end{aligned}$$

donde $p(x) = ax^2 + bx + c$ es un polinomio cuadrático con $a = \mathbf{u} \cdot \mathbf{u}$, $b = 2\mathbf{u} \cdot \mathbf{v}$ y $c = \mathbf{v} \cdot \mathbf{v}$. Dado que tanto $p(x)$ como a son mayores o iguales a 0, la gráfica del polinomio es una parábola cóncava hacía arriba con vértice en el semiplano superior (por encima del eje X). Recordando que el vértice de $p(x)$ es $\left(-\frac{b}{2a}, c - \frac{b^2}{4a}\right)$, tenemos

$$0 \leq c - \frac{b^2}{4a} = \|\mathbf{v}\|^2 - \frac{4(\mathbf{u} \cdot \mathbf{v})^2}{4\|\mathbf{u}\|^2},$$

de donde obtenemos que

$$(\mathbf{u} \cdot \mathbf{v})^2 \leq \|\mathbf{u}\|^2 \ \|\mathbf{v}\|^2$$

y tomando raíz cuadrada, concluimos que

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \ \|\mathbf{v}\|.$$

Además, si $\mathbf{u} = \lambda \mathbf{v}$,

$$\begin{aligned} |\mathbf{u} \cdot \mathbf{v}| &= |\lambda \mathbf{v} \cdot \mathbf{v}| = |\lambda||\mathbf{v} \cdot \mathbf{v}| \\ &= |\lambda|\|\mathbf{v}\|^2 = |\lambda|\|\mathbf{v}\|\|\mathbf{v}\| \\ &= \|\lambda \mathbf{v}\|\|\mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\| \end{aligned}$$

□

Teorema 12 [*Desigualdad triangular*].
Dados los vectores $\mathbf{u}$ y $\mathbf{v}$ de R^n, tenemos que

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| \ + \|\mathbf{v}\|.$$

La igualdad se cumple, si y solo si, $\mathbf{u} = \lambda \mathbf{v}$ con $\lambda \geq 0$.

Demostración:

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \|\mathbf{u}\|^2 \ + \|\mathbf{v}\|^2 + 2(\mathbf{u} \cdot \mathbf{v}) \quad (\text{ Resultado 1 del Teorema 10}) \\ &\leq \|\mathbf{u}\|^2 \ + \|\mathbf{v}\|^2 + 2(\|\mathbf{u}\|\|\mathbf{v}\|) \ (\text{ Teorema 11}) \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2. \end{aligned}$$

Por tanto,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| \ + \|\mathbf{v}\|.$$

Además, si $\mathbf{u} = \lambda \mathbf{v}$, con $\lambda \geq 0$

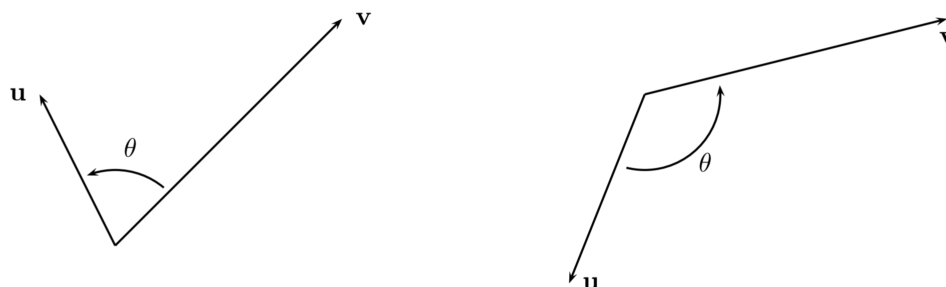
$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\| &= \|\lambda \mathbf{v} + \mathbf{v}\| = \|(\lambda + 1)\mathbf{v}\| \\ &= |(\lambda + 1)|\|\mathbf{v}\| = (\lambda + 1)\|\mathbf{v}\| \\ &= \lambda\|\mathbf{v}\| + \|\mathbf{v}\| = \|\lambda \mathbf{v}\| + \|\mathbf{v}\| \\ &= \|\mathbf{u}\| + \|\mathbf{v}\|. \end{aligned}$$

□

Veamos ahora que el producto escalar también nos permite calcular la medida del ángulo entre dos vectores no nulos de R^n, para lo cual, después de definir ángulo entre vectores, aplicamos el Teorema del Coseno.

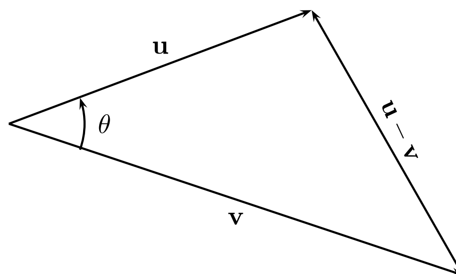
Definición 13 [*Ángulo entre vectores*]. Dados los vectores no nulos $\mathbf{u}$ y $\mathbf{v}$ de R^n, definimos el ángulo determinado por $\mathbf{u}$ y $\mathbf{v}$ como el menor giro positivo[6] que hace uno de ellos para coincidir con la dirección y sentido del otro, como vemos en la figura siguiente

[6] Entendemos por giro positivo aquel que se hace en el sentido contrario al de las manecillas del reloj.



Notemos que, si los vectores **u** y **v** son vectores libres, es posible mover paralelamente uno de ellos para que sus puntos iniciales coincidan, por lo cual, entre vectores que no tienen un mismo origen, también existe un ángulo.

Dados dos vectores **u** y **v** no nulos de R^n, siempre podemos construir un triángulo como el de la figura siguiente.



Al aplicar el Teorema del Coseno a este triángulo, tenemos

$$\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta.$$

De otro lado, utilizando el Teorema 10, obtenemos

$$\|\mathbf{u}\|^2 - 2\mathbf{u}\cdot\mathbf{v} + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta,$$

de donde, concluimos que $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta$. Esta igualdad nos brinda una forma para calcular la medida del ángulo entre dos vectores no nulos de R^n, tomando a θ como el ángulo tal que

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}. \tag{2.1}$$

Ejemplo 23. Calculemos el ángulo entre los vectores $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$.

Empecemos calculando $\mathbf{u}\cdot\mathbf{v}$, $\|\mathbf{u}\|^2$ y $\|\mathbf{v}\|^2$.

$$\mathbf{u}\cdot\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = 2, \quad \|\mathbf{u}\|^2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = 4 \quad \text{y} \quad \|\mathbf{v}\|^2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = 4.$$

Asi que

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{2}{\sqrt{4}\sqrt{4}} = \frac{1}{2};$$

por lo tanto, el ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$ es $\pi/3$, ya que $\pi/3$ es el menor ángulo tal que $\cos(\pi/3) = 1/2$.

□

De la fórmula (2.1) para calcular el ángulo entre dos vectores, tenemos que $\cos\theta$ y $\mathbf{u}\cdot\mathbf{v}$ tienen el mismo signo. Así que, como $\cos\theta$ es positivo cuando $0 \leq \theta < \pi/2$, negativo cuando $\pi/2 < \theta \leq \pi$ y cero cuando $\theta = \pi/2$, entonces

$\mathbf{u}\cdot\mathbf{v} > 0$	si y solo si	$0 \leq \theta < \pi/2$ (θ es un ángulo agudo)
$\mathbf{u}\cdot\mathbf{v} < 0$	si y solo si	$\pi/2 < \theta \leq \pi$ (θ es un ángulo obtuso)
$\mathbf{u}\cdot\mathbf{v} = 0$	si y solo si	$\theta = \pi/2$ (θ es un ángulo recto)

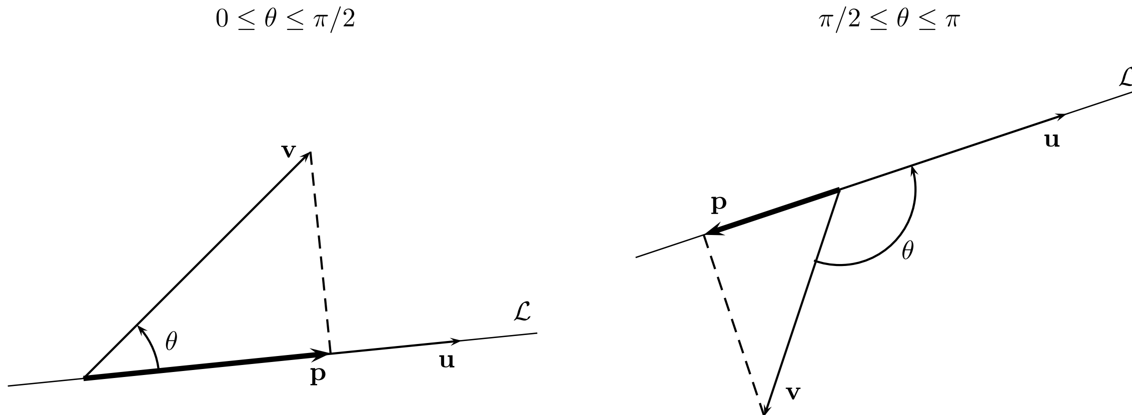
Como muy bien sabemos, el concepto de ortogonalidad es fundamental en la geometría de R^2 y R^3. Con las observaciones anteriores, podemos generalizar este concepto a vectores de R^n.

Definición 14 [*Vectores ortogonales*][7]. Diremos que los vectores no nulos $\mathbf{u}$ y $\mathbf{v}$ de R^n son ortogonales, si y solo si, $\mathbf{u}\cdot\mathbf{v} = 0$.

Ejemplo 24. Dados los vectores $\mathbf{a} = \begin{pmatrix} -1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 10 \end{pmatrix}$ y $\mathbf{c} = \begin{pmatrix} 0 \\ -1 \\ 5 \\ 4 \end{pmatrix}$, determinemos la(s) pareja(s) de vectores ortogonales.

Como $\mathbf{a}\cdot\mathbf{b} = 0$, el vector $\mathbf{a}$ es ortogonal al vector $\mathbf{b}$. De la misma forma, como $\mathbf{a}\cdot\mathbf{c} \neq 0$ y $\mathbf{b}\cdot\mathbf{c} \neq 0$, el vector $\mathbf{a}$ no es ortogonal al vector $\mathbf{c}$, ni el vector $\mathbf{b}$ es ortogonal al vector $\mathbf{c}$. □

Un tema importante, donde está presente el concepto de ortogonalidad, es la proyección ortogonal de un vector sobre otro, como lo describimos a continuación. Dados dos vectores no nulos $\mathbf{u}$ y $\mathbf{v}$ con un mismo punto inicial[8], llamemos $\mathbf{p}$ al vector que obtenemos al trazar una perpendicular desde el punto final del vector $\mathbf{v}$ sobre la recta $\mathcal{L}$ que contiene al vector $\mathbf{u}$ y θ al ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$, como vemos en la figura siguiente.



Aplicando un resultado básico de geometría (en un triángulo rectángulo, la magnitud de un cateto es el producto de la magnitud de la hipotenusa por el coseno del ángulo adyacente), tenemos que $\|\mathbf{p}\| = \|\mathbf{v}\||\cos\theta|$. De otro lado, es claro que, cuando $0 \leq \theta \leq \pi/2$, $\mathbf{p} = \|\mathbf{p}\|\tilde{\mathbf{u}}$; y cuando $\pi/2 < \theta \leq \pi$, $\mathbf{p} = -\|\mathbf{p}\|\tilde{\mathbf{u}}$, donde $\tilde{\mathbf{u}}$ es el vector unitario en la dirección y sentido de $\mathbf{u}$, es decir, $\tilde{\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$. Usando estas igualdades, cuando $0 \leq \theta \leq \pi/2$, obtenemos

$$\mathbf{p} = \|\mathbf{p}\|\ \tilde{\mathbf{u}} = \|\mathbf{v}\|\ |\cos\theta|\ \tilde{\mathbf{u}} = \|\mathbf{v}\|\ \cos\theta\ \tilde{\mathbf{u}}$$

[7]En la literatura, el término perpendicular es usado como sinónimo de ortogonal

[8]Si no tienen un mismo punto inicial, movemos paralelamente uno de ellos o ambos, hasta que sus puntos iniciales coincidan.

y cuando $\pi/2 < \theta \leq \pi$, obtenemos

$$\mathbf{p} = -\|\mathbf{p}\| \ \tilde{\mathbf{u}} = -\|\mathbf{v}\| \ |\cos\theta| \ \tilde{\mathbf{u}} = -\|\mathbf{v}\| \ (-\cos\theta) \ \tilde{\mathbf{u}} = \|\mathbf{v}\| \ \cos\theta \ \tilde{\mathbf{u}}.$$

En ambos casos, usando (2.1),

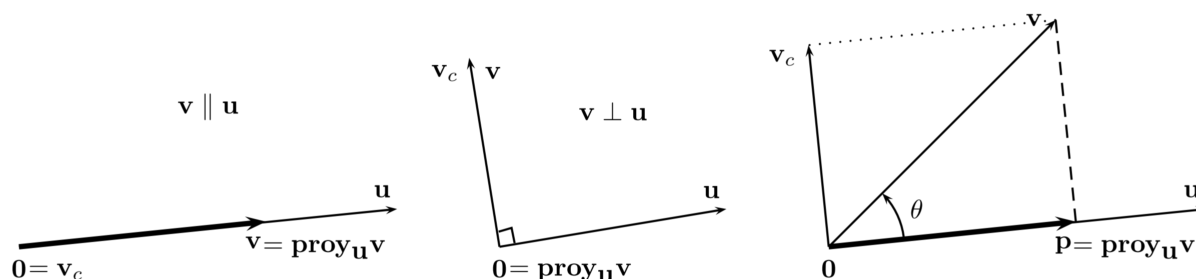
$$\begin{aligned} \mathbf{p} &= \|\mathbf{v}\| \cos\theta \ \frac{\mathbf{u}}{\|\mathbf{u}\|} \\ &= \|\mathbf{v}\| \ \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \ \frac{\mathbf{u}}{\|\mathbf{u}\|} \\ &= \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\right) \ \mathbf{u} \ . \end{aligned}$$

Con base en este resultado, definimos el concepto de proyección ortogonal para vectores de R^n.

Definición 15 [*Proyección ortogonal*]. Si $\mathbf{u} \neq \mathbf{0}$ y $\mathbf{v}$ son vectores de R^n, definimos $\mathbf{proy_u v}$, la *proyección ortogonal* de $\mathbf{v}$ sobre $\mathbf{u}$, como el vector

$$\mathbf{proy_u v} = \left(\frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{u}\|^2}\right) \ \mathbf{u}.$$

Si llamamos $\mathbf{v}_c = \mathbf{v} - \mathbf{proy_u v}$, tenemos que, si $\mathbf{u}$ y $\mathbf{v}$ son ortogonales, $\mathbf{proy_u v} = \mathbf{0}$ y por tanto, $\mathbf{v}_c = \mathbf{v}$; si $\mathbf{u}$ y $\mathbf{v}$ son paralelos, $\mathbf{proy_u v} = \mathbf{v}$ y por tanto, $\mathbf{v}_c = \mathbf{0}$; y, en los demás casos, $\mathbf{v}_c$ y $\mathbf{u}$ son vectores ortogonales y por tanto, $\mathbf{v}_c$ y $\mathbf{proy_u v}$ también lo son. Así, si $\mathbf{u}$ y $\mathbf{v}$ no son ortogonales ni paralelos, podemos expresar el vector $\mathbf{v}$ como la suma de dos vectores ortogonales, $\mathbf{proy_u v}$ y $\mathbf{v}_c$, lo que sugiere llamar a $\mathbf{v}_c$ la *componente vectorial de* $\mathbf{v}$ *ortogonal a* $\mathbf{u}$.



Ejemplo 25. Encontremos la proyección del vector $\mathbf{v}$ sobre el vector $\mathbf{u}$ y la componente vectorial de $\mathbf{v}$ ortogonal a $\mathbf{u}$, para cada uno de los siguientes casos, en los que ella esté definida.

(a) $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, **(b)** $\mathbf{u} = \mathbf{e}_1$ y $\mathbf{v} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix}$, **(c)** $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -2 \end{pmatrix}$ y $\mathbf{u} = \begin{pmatrix} 3 \\ -3 \\ -2 \\ 1 \end{pmatrix}$,

(d) $\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} -3 \\ -3 \\ 6 \end{pmatrix}$.

(a) Calculemos $\mathbf{u}\cdot\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = 1$ y $\|\mathbf{u}\|^2 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = 3$. Así que

$$\mathbf{proy_u v} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\right) \ \mathbf{u} = \frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -1/3 \\ -1/3 \end{pmatrix}$$

y

$$\mathbf{v}_c = \mathbf{v} - \mathbf{proy}_{\mathbf{u}}\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/3 \\ -1/3 \\ -1/3 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -2/3 \\ 4/3 \end{pmatrix}.$$

(b) Puesto que el vector $\mathbf{e}_1$ tiene norma 1 y $\mathbf{v} \cdot \mathbf{e}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 2$, tenemos que

$$\mathbf{proy}_{\mathbf{u}}\mathbf{v} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}\right) \mathbf{u} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y

$$\mathbf{v}_c = \mathbf{v} - \mathbf{proy}_{\mathbf{u}}\mathbf{v} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 3 \end{pmatrix}.$$

(c) Veamos que $\mathbf{u} \cdot \mathbf{v} = \begin{pmatrix} 3 \\ -3 \\ -2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \\ -2 \end{pmatrix} = 0$ (es decir, $\mathbf{u}$ y $\mathbf{v}$ son ortogonales) y que por tanto

$$\mathbf{proy}_{\mathbf{u}}\mathbf{v} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}\right) \mathbf{u} = \frac{0}{\|\mathbf{u}\|^2} \begin{pmatrix} 3 \\ -3 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

(d) Tenemos que $\mathbf{v} = \begin{pmatrix} -3 \\ -3 \\ 6 \end{pmatrix} = -3 \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = -3\mathbf{u}$ ($\mathbf{u}$ y $\mathbf{v}$ son paralelos), $\begin{pmatrix} -3 \\ -3 \\ 6 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = -18$ y que $\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = 6$, entonces

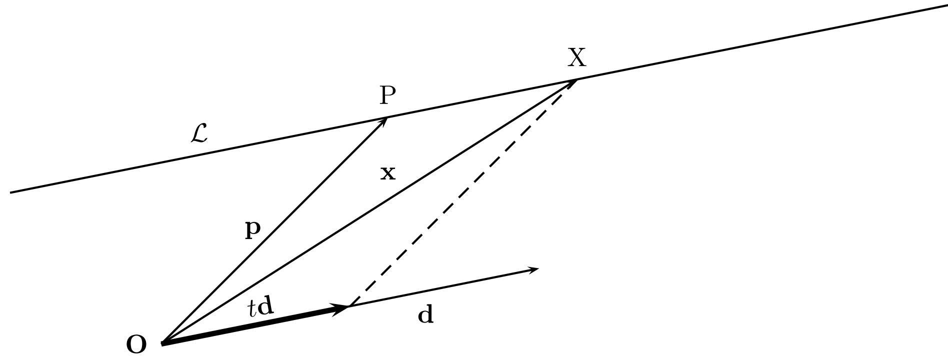
$$\mathbf{proy}_{\mathbf{u}}\mathbf{v} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}\right) \mathbf{u} = \frac{-18}{6} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = -3 \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \mathbf{v}.$$

□

2.8. Rectas, Planos e Hiperplanos

Aunque estamos familiarizados con la ecuación de la recta en el plano cartesiano (ecuación lineal en dos variables), en esta sección, con base en los conceptos, operaciones y propiedades de los vectores hasta aquí estudiados, presentaremos la ecuación de la recta desde un punto de vista vectorial y, en este mismo sentido, estudiaremos las ecuaciones de los planos y los hiperplanos en R^n. Comencemos con caracterizar los puntos de una recta en R^n.

Definición 16 [*Recta*]. Dado un punto $P \in R^n$ y un vector $\mathbf{d}$ no nulo de R^n, diremos que la recta $\mathcal{L}$ que contiene a P y tiene dirección $\mathbf{d}$ es el conjunto formado por P y todos los puntos X que determinan vectores $\overline{PX}$ paralelos a $\mathbf{d}$. Al vector $\mathbf{d}$ lo llamamos *vector dirección* o *vector director de la recta*.



De la definición anterior, tenemos que dada una recta $\mathcal{L}$, si P es un punto de la recta, $\mathbf{d}$ su vector director y X un punto cualquiera de la recta, el segmento dirigido o vector libre $\overline{PX}$ es paralelo a $\mathbf{d}$; es decir, es un múltiplo por escalar de $\mathbf{d}$. Como $\overline{PX}$ es la diferencia de los vectores $\mathbf{x} = \overline{OX}$ y $\mathbf{p} = \overline{OP}$, tenemos que para algún $t \in R$,

$$\mathbf{x} - \mathbf{p} = t\mathbf{d},$$

de donde,

$$\mathbf{x} = \mathbf{p} + t\mathbf{d}. \tag{2.2}$$

A la expresión (2.2) la llamamos *ecuación vectorial de la recta*, la cual también puede expresarse en término de sus componentes. Supongamos que el punto dado de la recta P, el vector director $\mathbf{d}$ y un punto arbitrario X de la recta, tienen componentes $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, $\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$ y $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, respectivamente. Entonces, teniendo en cuenta que $\mathbf{x} = \overline{OX}$, podemos escribir la ecuación (2.2) como

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + t \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$$

o, lo que es equivalente,

$$\begin{array}{rcl} x_1 & = & a_1 + td_1 \\ x_2 & = & a_2 + td_2 \\ & \vdots & \\ x_n & = & a_n + td_n. \end{array} \tag{2.3}$$

A estas últimas ecuaciones las llamamos *ecuaciones paramétricas de la recta*.

Ejemplo 26. Dada la ecuación vectorial de la recta $\mathcal{L}$: $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix}$.

(a) Encontremos dos puntos P y Q de la recta $\mathcal{L}$.

(b) Determinemos si los puntos $R = \begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix}$ y $S = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix}$ pertenecen a la recta $\mathcal{L}$.

(c) Encontremos un vector $\mathbf{d}$ que sea un vector director de la recta $\mathcal{L}$.

(d) Verifiquemos que el vector $\overline{PQ}$, donde P y Q son los puntos hallados en (a), es paralelo a $\mathbf{d}$, el vector director de la recta $\mathcal{L}$ encontrado en **(c)**.

(a) Para encontrar puntos de la recta $\mathcal{L}$, es suficiente que le demos valores arbitrarios al parámetro t. Por ejemplo, si $t = 1$ y luego $t = -2$ en la ecuación vectorial de la recta, obtenemos dos puntos P y Q de la recta $\mathcal{L}$:

$$\begin{aligned} P &= \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + 1 \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix} &= \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix} \\ Q &= \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + (-2) \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix} &= \begin{pmatrix} 4 \\ -1 \\ -7 \end{pmatrix}. \end{aligned}$$

De esta forma, podemos hallar tantos puntos de la recta $\mathcal{L}$ como queramos.

(b) Para determinar si $R = \begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix}$ se encuentra en la recta $\mathcal{L}$, tenemos que verificar si existe un escalar t tal que

$$\begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix}, \tag{2.4}$$

o lo que es lo mismo, verificar si existe un escalar t tal que

$$\begin{aligned} 3 &= 2 - t \\ -1 &= -1 \\ -2 &= 3 + 5t. \end{aligned}$$

De la primera ecuación, tenemos que $t = -1$, y si sustituimos el valor de t por -1 en la segunda y tercera ecuación, obtenemos proposiciones verdaderas, lo que implica que al sustituir t por -1 en (2.4), obtenemos una proposición verdadera y por tanto, R es un punto de la recta $\mathcal{L}$.

Con el mismo razonamiento, para verificar si S es un punto de la recta $\mathcal{L}$, debemos verificar si existe t tal que

$$\begin{aligned} 4 &= 2 - t \\ -1 &= -1 \\ 0 &= 3 + 5t. \end{aligned}$$

De la primera ecuación, $t = -2$, pero al sustituir t por -2 en la tercera ecuación, tenemos $0 = -7$, una proposición falsa. Esto nos indica que S no es un punto de la recta.

(c) De la ecuación vectorial de la recta $\mathcal{L}$, un vector director de dicha recta es $\mathbf{d} = \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix}$ o cualquier vector no nulo paralelo a él.

(d) Sabemos que, en término de coordenadas, $\overline{PQ} = Q - P = \begin{pmatrix} 4 \\ -1 \\ -7 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -15 \end{pmatrix}$. Y, puesto que $\begin{pmatrix} 3 \\ 0 \\ -15 \end{pmatrix} = (-3) \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix}$, podemos concluir que efectivamente $\overline{PQ}$ es paralelo al vector director de la recta $\mathcal{L}$. □

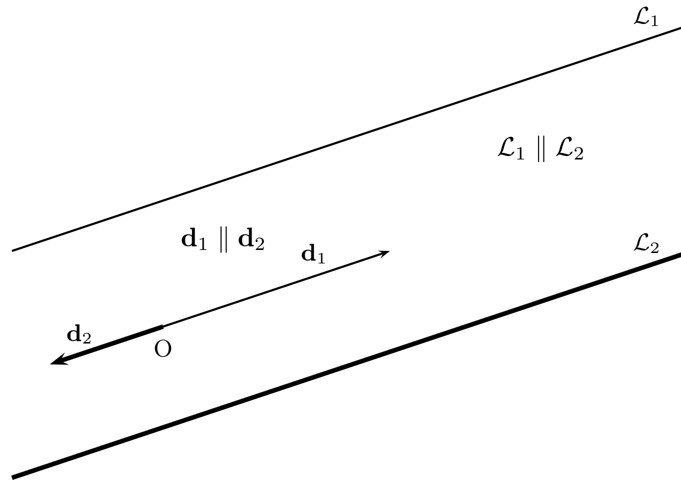
Ejemplo 27. Encontremos la ecuación vectorial de la recta que pasa por los puntos

$$P = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \text{ y } Q = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

Si la recta contiene a los puntos P y Q, ella tiene la dirección $\mathbf{d} = \overline{PQ} = Q - P = \begin{pmatrix} 3 \\ -1 \\ -1 \\ 0 \end{pmatrix}$. Por tanto, la ecuación vectorial de la recta es $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 3 \\ -1 \\ -1 \\ 0 \end{pmatrix}$. □

Es natural que la ecuación vectorial de una recta no sea única, puesto que existe un número infinito de puntos y de vectores que pueden ser seleccionados como P y $\mathbf{d}$, respectivamente. De otro lado, un mismo vector puede ser vector director de infinitas rectas, las cuales son un caso particular de las que llamaremos *rectas paralelas*.

Definición 17 [*Rectas paralelas*]. Sean $\mathcal{L}_1$ y $\mathcal{L}_2$ dos rectas en R^n, con vectores directores $\mathbf{d}_1$ y $\mathbf{d}_2$, respectivamente. Diremos que las rectas $\mathcal{L}_1$ y $\mathcal{L}_2$ son paralelas, si y solo si, los vectores $\mathbf{d}_1$ y $\mathbf{d}_2$ lo son.



Observemos que de acuerdo con el criterio dado de vectores paralelos, podemos decir que dos rectas son paralelas, si y solo si, el vector director de una de ellas es múltiplo por escalar del vector director de la otra.

Ejemplo 28. Determinemos si las siguientes rectas son paralelas.

(**a**) $\mathcal{L}_1$ es la recta que pasa por los puntos $P = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ y $Q = \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix}$ y $\mathcal{L}_2$ es la recta con ecuación vectorial dada por $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -4 \\ 3 \end{pmatrix} + t \begin{pmatrix} 4 \\ 10 \\ -2 \end{pmatrix}$.

(**b**) $\mathcal{L}_1$ es la recta que pasa por el punto $M = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ y tiene vector dirección $\mathbf{v} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ y $\mathcal{L}_2$ es la recta que pasa por los puntos $Q = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$ y $R = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$.

(**a**) Un vector director de la recta $\mathcal{L}_1$ es $\mathbf{d}_1 = \overline{PQ} = Q - P = \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix} - \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}$ y un vector

director de la recta $\mathcal{L}_2$ es $\mathbf{d}_2 = \begin{pmatrix} 4 \\ 10 \\ -2 \end{pmatrix}$. Como $\mathbf{d}_2 = 2\mathbf{d}_1$, tenemos que $\mathbf{d}_1$ y $\mathbf{d}_2$ son paralelos y por tanto, las rectas $\mathcal{L}_1$ y $\mathcal{L}_2$ también lo son.

(**b**) Tenemos que un vector director de la recta $\mathcal{L}_2$ es $\overline{QR} = R - Q = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix}$.

Es claro que no existe $\lambda \in R$, tal que $\begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$, así que los vectores directores de las rectas no son paralelos y por tanto, las rectas $\mathcal{L}_1$ y $\mathcal{L}_2$ tampoco lo son. $\square$

Una aplicación interesante de la anterior definición está dada por la caracterización de la igualdad de dos rectas, como lo plantea el siguiente teorema.

Teorema 13 [*Rectas iguales*].
Dos rectas son iguales, si y solo si, las dos rectas son paralelas y tienen al menos un punto en común.

Demostración: Sean $\mathbf{d}_1$ y $\mathbf{d}_2$ vectores directores de las rectas $\mathcal{L}_1$ y $\mathcal{L}_2$, respectivamente.

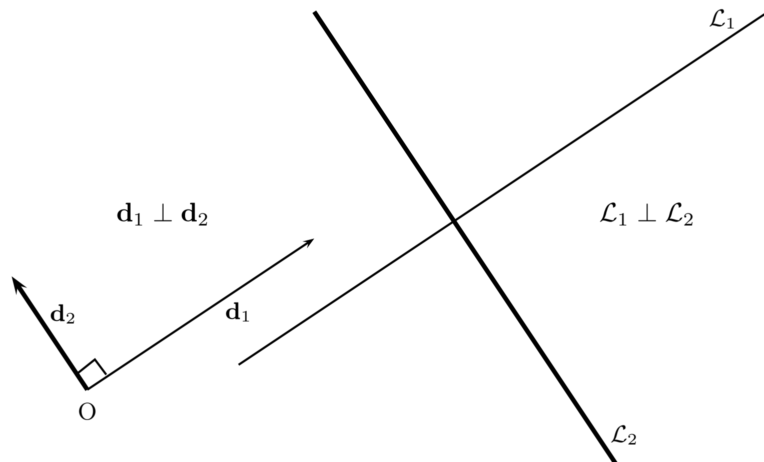
Si $\mathcal{L}_1 = \mathcal{L}_2$, tomemos dos puntos $P,\ Q \in \mathcal{L}_1 = \mathcal{L}_2,\ P \neq Q$. Por definición de recta, existen escalares α y β diferentes de cero tales que $\overline{PQ} = \alpha\mathbf{d}_1$ y $\overline{PQ} = \beta\mathbf{d}_2$, de donde $\alpha\mathbf{d}_1 = \beta\mathbf{d}_2$; por lo tanto, $\mathbf{d}_1$ es paralelo a $\mathbf{d}_2$, y por la definición anterior, $\mathcal{L}_1$ es paralela a $\mathcal{L}_2$.

De otro lado, sea $P \in \mathcal{L}_1 \cap \mathcal{L}_2$. Si $\mathcal{L}_1$ es paralela a $\mathcal{L}_2$, por la definición de rectas paralelas, $\mathbf{d}_1$ es paralelo a $\mathbf{d}_2$; es decir, existe $\lambda \neq 0$ tal que $\mathbf{d}_1 = \lambda\mathbf{d}_2$. Demostremos que cualquier punto de la recta $\mathcal{L}_1$ pertenece también a la recta $\mathcal{L}_2$ y viceversa.

Si $X \in \mathcal{L}_1$, por la definición de recta, $\overline{PX} = \alpha\mathbf{d}_1$ para algún $\alpha \in R$; por lo tanto, $\overline{PX} = \alpha(\lambda\mathbf{d}_2) = (\alpha\lambda)\mathbf{d}_2$; es decir, existe $\beta = \alpha\lambda \in R$ tal que $\overline{PX} = \beta\mathbf{d}_2$; por lo tanto, $X \in \mathcal{L}_2$.
Similarmente, Si $X \in \mathcal{L}_2$, por la definición de recta, $\overline{PX} = \rho\mathbf{d}_2$ para algún $\rho \in R$; por lo tanto, $\overline{PX} = \rho(\frac{1}{\lambda}\mathbf{d}_1) = (\frac{\rho}{\lambda})\mathbf{d}_1$; es decir, existe $\delta = \frac{\rho}{\lambda} \in R$ tal que $\overline{PX} = \delta\mathbf{d}_1$; por lo tanto, $X \in \mathcal{L}_1$. $\square$

Al igual que el concepto de rectas paralelas, otro concepto importante en la geometría Euclidiana es el de rectas ortogonales.

Definición 18 [*Rectas ortogonales*]. Sean $\mathcal{L}_1$ y $\mathcal{L}_2$ dos rectas con vectores directores $\mathbf{d}_1$ y $\mathbf{d}_2$, respectivamente. Diremos que las rectas $\mathcal{L}_1$ y $\mathcal{L}_2$ son ortogonales, si y solo si, los vectores $\mathbf{d}_1$ y $\mathbf{d}_2$ lo son; es decir, si y solo si, $\mathbf{d}_1 \cdot \mathbf{d}_2 = 0$



Ejemplo 29. Determinemos si las siguientes rectas son ortogonales.

(**a**) $\mathcal{L}_1$ es la recta que pasa por los puntos $P = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ y $Q = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$ y $\mathcal{L}_2$ es la recta con ecuación vectorial dada por $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \\ 1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix}$.

(**b**) $\mathcal{L}_1$ es la recta que pasa por el punto $M = \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}$ y tiene vector dirección $\mathbf{v} = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$ y $\mathcal{L}_2$ es la recta que pasa por los puntos $Q = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ y $R = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$.

(**a**) Un vector director de la recta $\mathcal{L}_1$ es $\mathbf{d}_1 = \overline{PQ} = Q - P = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$ y un vector director de la recta $\mathcal{L}_2$ es $\mathbf{d}_2 = \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix}$. Como $\mathbf{d}_1 \cdot \mathbf{d}_2 = 0$, tenemos que $\mathbf{d}_1$ y $\mathbf{d}_2$ son ortogonales y por tanto, las rectas $\mathcal{L}_1$ y $\mathcal{L}_2$ también lo son.

(**b**) Tenemos que un vector director de la recta $\mathcal{L}_2$ es $\mathbf{d}_2 = \overline{QR} = R - Q = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ -2 \end{pmatrix}$.
Como $\mathbf{d}_2 \cdot \mathbf{v} = 1 \cdot 1 + 3 \cdot 5 + (-1) \cdot (-2) = 18 \neq 0$, los vectores directores de las rectas no son ortogonales y por tanto, $\mathcal{L}_1$ y $\mathcal{L}_2$ tampoco lo son. □

Al analizar el sistema de ecuaciones (2.3), ecuaciones paramétricas de la recta, observamos que el parámetro t puede ser eliminado, al despejar t de cada ecuación e igualar esta expresiones, ya que el vector director $\mathbf{d}$ debe ser diferente de cero. Así, reducimos el sistema a otro con $n-1$ ecuaciones lineales y n incógnitas, donde las incógnitas son las componentes de los puntos que conforman la recta. A estas ecuaciones lineales, que podemos escribir como

$$\frac{x_1 - a_1}{d_1} = \frac{x_2 - a_2}{d_2} = \ldots = \frac{x_n - a_n}{d_n} \quad \text{siempre que} \quad d_i \neq 0,\ i = 1, 2, \ldots n,$$

las llamamos *ecuaciones simétricas de la recta* que pasa por el punto P y tiene vector director $\mathbf{d}$, y, por su simplicidad, son de especial utilidad en R^3. Es de anotar que, cuando $d_i = 0$ para algún i, a cambio de la fracción $\frac{x_i - a_i}{d_i}$ en la anterior expresión, incluimos la ecuación $x_i = a_i$, como mostramos en los siguientes ejemplos.

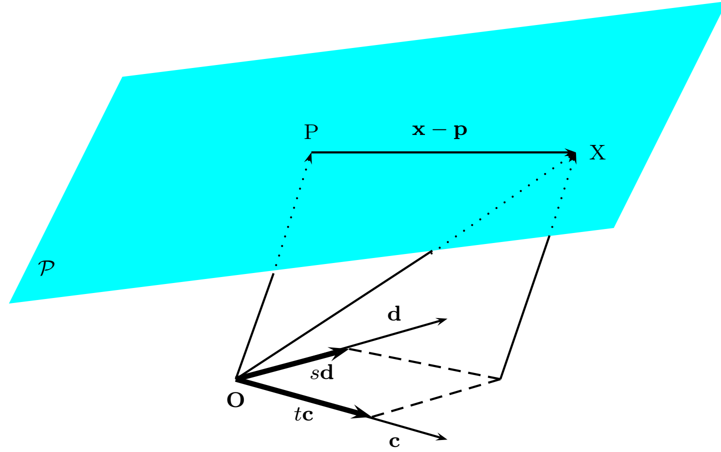
Ejemplo 30. Escribamos las ecuaciones simétricas de las rectas $\mathcal{L}_2$ del Ejemplo 28 y de la recta hallada en el Ejemplo 27.

Las ecuaciones simétricas de las rectas $\mathcal{L}_2$ del Ejemplo 28 son:
a. $\frac{x-0}{4} = \frac{y-(-4)}{10} = \frac{z-3}{-2}$, las cuales también podemos escribir como $\frac{x}{4} = \frac{y+4}{10} = \frac{3-z}{2}$.
b. $\frac{x-2}{2} = \frac{y-3}{5}$, $z = 1$, las cuales también podemos escribir como $5(x-2) = 2(y-3)$, $z = 1$.

Las ecuaciones simétricas de la recta encontrada en el Ejemplo 27 son $\frac{x_1-(-1)}{3} = \frac{x_2-2}{-1} = \frac{x_3-0}{-1}$, $x_4 = 1$, las cuales también podemos escribir como $\frac{x_1+1}{3} = 2 - x_2 = -x_3$, $x_4 = 1$. □

De la misma forma como, dado un punto P, al conjunto de puntos X que determinan vectores $\overline{PX}$ que resultan ser múltiplos por escalar de un vector dado $\mathbf{d}$, lo llamamos recta, al conjunto de puntos X que determinan vectores $\overline{PX}$ que resultan ser combinación lineal de dos vectores dados $\mathbf{c}$ y $\mathbf{d}$, lo llamamos de una manera especial: *Plano.*

Definición 19 [*Plano*]. Dado un punto $P \in R^n$ y dos vectores $\mathbf{c}$ y $\mathbf{d} \in R^n$ diferentes de cero y no paralelos, diremos que el conjunto de puntos X que determinan vectores $\overline{PX}$ que son combinación lineal de los vectores $\mathbf{c}$ y $\mathbf{d}$, es el *plano* $\mathcal{P}$ que pasa por el punto P y tiene direcciones $\mathbf{c}$ y $\mathbf{d}$. A los vectores $\mathbf{c}$ y $\mathbf{d}$ los llamamos *vectores directores* o *vectores dirección del plano*.



De la definición anterior, tenemos que dado un plano $\mathcal{P}$, si P es un punto del plano, $\mathbf{c}$ y $\mathbf{d}$ son dos vectores directores del plano y X es un punto cualquiera del plano, entonces $\overline{PX} = t\mathbf{c} + s\mathbf{d}$ con $t, s \in R$. Así, si llamamos $\mathbf{x}$ al vector $\overline{OX}$ y $\mathbf{p}$ al vector $\overline{OP}$, tenemos que, para $t, s \in R$,

$$\mathbf{x} - \mathbf{p} = t\mathbf{c} + s\mathbf{d},$$

lo que es equivalente a

$$\mathbf{x} = \mathbf{p} + t\mathbf{c} + s\mathbf{d},$$

la cual llamamos *ecuación vectorial del plano*. Al expresar esta última ecuación en término de las componentes de los vectores $\mathbf{p} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$, $\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$ y $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, obtenemos

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + t \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} + s \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix},$$

de donde, tenemos el sistema de ecuaciones

$$\begin{aligned} x_1 &= a_1 + tc_1 + sd_1 \\ x_2 &= a_2 + tc_2 + sd_2 \\ &\vdots \\ x_n &= a_n + tc_n + sd_n. \end{aligned}$$

A estas ecuaciones las llamamos *ecuaciones paramétricas del plano*.

Ejemplo 31. Dadas las ecuaciones paramétricas del plano $\mathcal{P}$

$$\begin{aligned} x_1 &= 2+t-s \\ x_2 &= 2t \\ x_3 &= 1+5s \\ x_4 &= -2 \end{aligned}$$

(a) Encontremos tres puntos P, Q y R del plano $\mathcal{P}$.
(b) Encontremos dos vectores $\mathbf{c}$ y $\mathbf{d}$ que sean vectores directores del plano $\mathcal{P}$.
(c) Verifiquemos que los vectores $\overline{PQ}$, $\overline{PR}$ y $\overline{QR}$ son combinaciones lineales de $\mathbf{c}$ y $\mathbf{d}$, los vectores directores del plano $\mathcal{P}$ encontrados en **(b)**.
(d) Determinemos si los puntos $M = \begin{pmatrix} 2 \\ 2 \\ 1 \\ -2 \end{pmatrix}$ y $N = \begin{pmatrix} 6 \\ 4 \\ -9 \\ -2 \end{pmatrix}$ se encuentran en el plano $\mathcal{P}$.

(a) Para encontrar puntos del plano $\mathcal{P}$, demos valores arbitrarios a los parámetros t y s. Por ejemplo, si $t = s = 0$ en las ecuaciones paramétricas del plano, tenemos que $x_1 = 2$, $x_2 = 0$, $x_3 = 1$ y $x_4 = -2$, por lo tanto, el punto $P = \begin{pmatrix} 2 \\ 0 \\ 1 \\ -2 \end{pmatrix} \in \mathcal{P}$. Similarmente, si $t = 1$ y $s = 0$, obtenemos que el punto $Q = \begin{pmatrix} 3 \\ 2 \\ 1 \\ -2 \end{pmatrix} \in \mathcal{P}$.

Finalmente, si $t = 1$ y $s = -1$, obtenemos que $R = \begin{pmatrix} 4 \\ 2 \\ -4 \\ -2 \end{pmatrix} \in \mathcal{P}$. De esta forma, podemos hallar tantos puntos del plano $\mathcal{P}$ como queramos.

(b) La ecuación vectorial del plano $\mathcal{P}$ es

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \\ -2 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -1 \\ 0 \\ 5 \\ 0 \end{pmatrix};$$

por lo tanto, dos vectores directores del plano son $\mathbf{c} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ y $\mathbf{d} = \begin{pmatrix} -1 \\ 0 \\ 5 \\ 0 \end{pmatrix}$.

(c) De la definición de vector libre, tenemos que $\overline{PQ} = Q - P = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}$, $\overline{PR} = R - P = \begin{pmatrix} 2 \\ 2 \\ -5 \\ 0 \end{pmatrix}$ y $\overline{QR} = R - Q = \begin{pmatrix} 1 \\ 0 \\ -5 \\ 0 \end{pmatrix}$. Ahora, veamos que estos vectores son combinación lineal de los vectores $\mathbf{c}$ y $\mathbf{d}$ hallados en **(b)**; es decir, que los sistemas de ecuaciones lineales, cuyas matriz aumentada conjunta es

$$\left(\begin{array}{rr|rrr} 1 & -1 & 1 & 2 & 1 \\ 2 & 0 & 2 & 2 & 0 \\ 0 & 5 & 0 & -5 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

son consistentes. Efectivamente, al escalonar esta matriz, obtenemos

$$\left(\begin{array}{cc|ccc} 1 & -1 & 1 & 2 & 1 \\ 0 & 2 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right);$$

de donde, podemos concluir que los sistemas son consistentes, y por lo tanto, los vectores $\overline{PQ}$, $\overline{PR}$ y $\overline{QR}$ son combinaciones lineales de $\mathbf{c}$ y $\mathbf{d}$.

(d) Para determinar si M y N se encuentran en el plano $\mathcal{P}$, tenemos que verificar si existen escalares t, s, α y β tales que

$$M = P + t\mathbf{c} + s\mathbf{d} \quad y \quad N = P + \alpha\mathbf{c} + \beta\mathbf{d}$$

donde, P es el punto encontrado en **(a)** y $\mathbf{c}$ y $\mathbf{d}$ son los vectores encontrados en **(b)**. En otras palabras, veamos si $M - P = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ y $N - P = \begin{pmatrix} 4 \\ 4 \\ -10 \\ 0 \end{pmatrix}$ son combinaciones lineales de $\mathbf{c}$ y $\mathbf{d}$; lo cual es equivalente a verificar si los sistemas de ecuaciones lineales, cuyas matriz aumentada conjunta es

$$\left(\begin{array}{cc|cc} 1 & -1 & 0 & 4 \\ 2 & 0 & 2 & 4 \\ 0 & 5 & 0 & -10 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

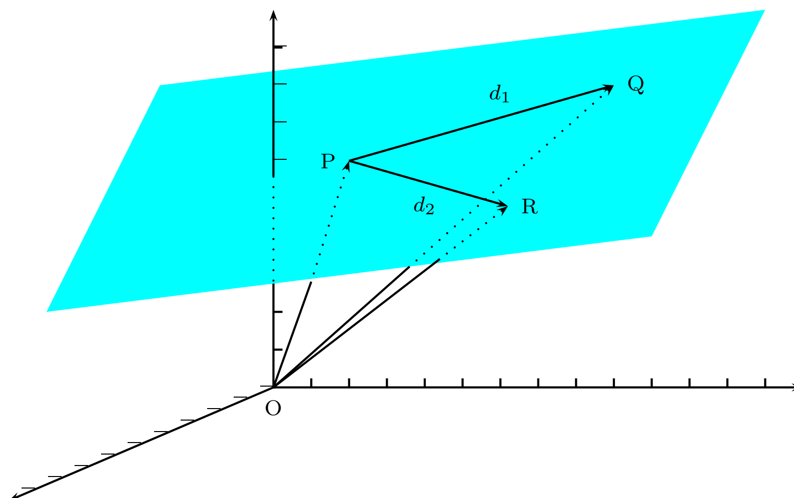
son consistentes. Al escalonar esta matriz, obtenemos

$$\left(\begin{array}{cc|cc} 1 & -1 & 0 & 4 \\ 0 & 2 & 2 & -4 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right);$$

de donde, podemos concluir que el primer sistema es inconsistente y que el segundo es consistente; y por lo tanto, que M no se encuentra en el plano $\mathcal{P}$, mientras que N si. □

Ejemplo 32. Encontremos una ecuación vectorial del plano que contiene los puntos $P = \begin{pmatrix} -2 \\ 5 \\ 3 \end{pmatrix}$, $Q = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$ y $R = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$.

El plano que contiene los puntos P, Q y R tiene como vectores directores a $\mathbf{d}_1 = \overline{PQ}$ y $\mathbf{d}_2 = \overline{PR}$, como lo muestra la siguiente gráfica.



Para el caso que nos ocupa, calculemos estos vectores.

$$\mathbf{d}_1 = \overline{PQ} = \begin{pmatrix} 2 \\ -7 \\ -2 \end{pmatrix} \text{ y } \mathbf{d}_2 = \overline{PR} = \begin{pmatrix} 4 \\ -5 \\ -6 \end{pmatrix}.$$

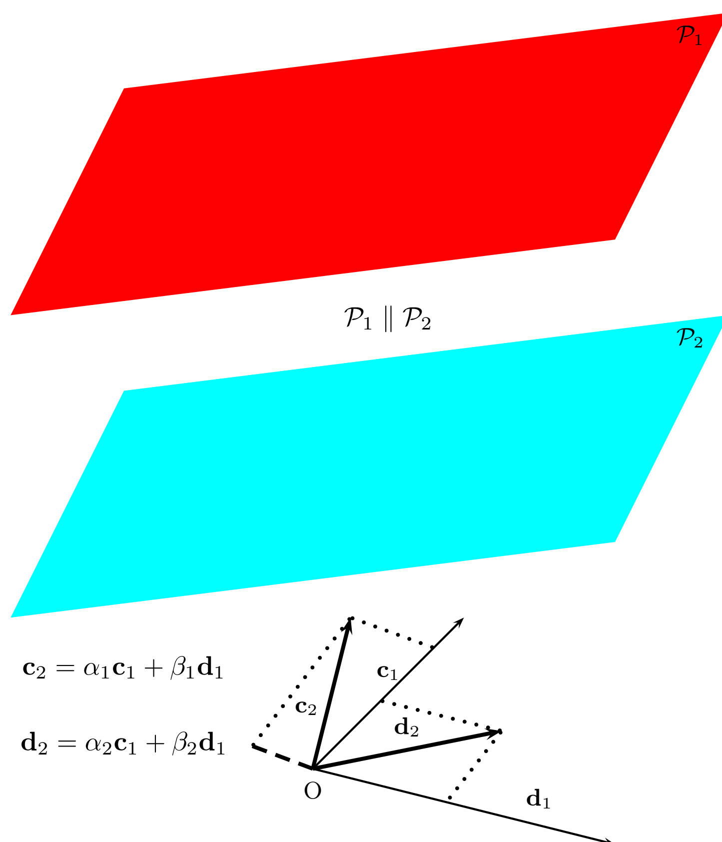
Asi, que una ecuación vectorial del plano que pasa por estos puntos es

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 5 \\ 3 \end{pmatrix} + t \begin{pmatrix} 2 \\ -7 \\ -2 \end{pmatrix} + s \begin{pmatrix} 4 \\ -5 \\ -6 \end{pmatrix}, \quad t,\ s \in R.$$

□

Al igual que para las rectas, es de esperarse que la ecuación de un plano no sea única; por ello, la importancia de establecer criterios claros para determinar cuando dos ecuaciones corresponden a un mismo plano. Empecemos por precisar cuando dos planos son paralelos.

Definición 20 [*Planos paralelos*]. Diremos que dos planos son paralelos, si y solo si, los vectores directores de uno de los planos son combinación lineal de los vectores directores del otro plano.



Ejemplo 33. Determinemos si el plano $\mathcal{P}_1$ que contiene los puntos $P = \begin{pmatrix} -2 \\ 5 \\ 3 \end{pmatrix}$, $Q = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$ y $R = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$ es paralelo al plano $\mathcal{P}_2$ que contiene al punto $P = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ y tiene vectores directores $\mathbf{c}_2 = \begin{pmatrix} 0 \\ 9 \\ -2 \end{pmatrix}$ y $\mathbf{d}_2 = \begin{pmatrix} -2 \\ -2 \\ 4 \end{pmatrix}$.

Del Ejemplo 32, tenemos que dos vectores directores del plano que contiene los puntos P, Q y R son

$$\mathbf{c}_1 = \overline{PQ} = \begin{pmatrix} 2 \\ -7 \\ -2 \end{pmatrix} \quad \text{y} \quad \mathbf{d}_1 = \overline{PR} = \begin{pmatrix} 4 \\ -5 \\ -6 \end{pmatrix}.$$

Ahora, para determinar si los planos señalados son paralelos, tenemos que chequear si existen escalares α, β, λ y μ, tales que

$$\begin{pmatrix} 2 \\ -7 \\ -2 \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 9 \\ -2 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ -2 \\ 4 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 4 \\ -5 \\ -6 \end{pmatrix} = \lambda \begin{pmatrix} 0 \\ 9 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ -2 \\ 4 \end{pmatrix}. \tag{2.5}$$

Al escalonar las matriz aumentada conjunta correspondiente a los sistemas anteriores, obtenemos

$$\left(\begin{array}{cc|cc} -2 & 4 & -2 & -6 \\ 0 & 16 & -16 & -32 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

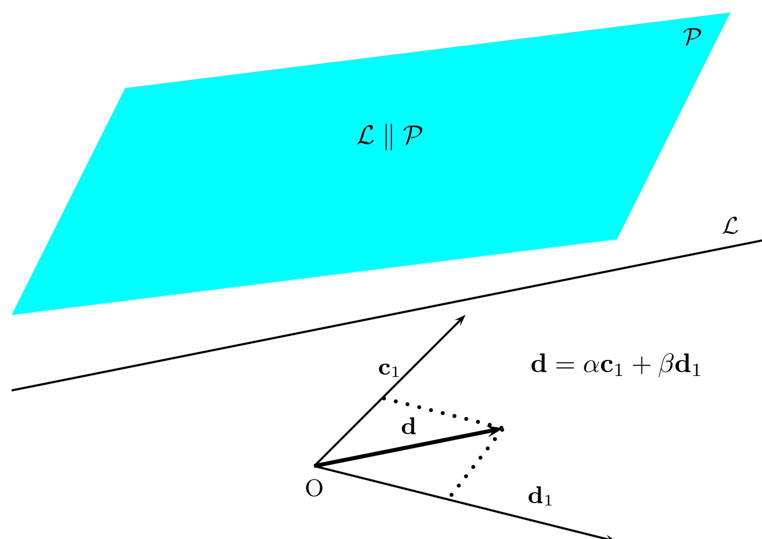
lo que implica que ambos sistemas son consistentes; es decir, que efectivamente existen escalares α, β, λ y μ, que satisfacen (2.5), por lo tanto, los planos son paralelos. □

Al igual que con el concepto de rectas paralelas, una aplicación interesante del concepto anterior es una caracterización de planos iguales, como lo expresa el siguiente teorema, cuya demostración la dejamos como ejercicio para el lector.

Teorema 14 [*Planos iguales*].
Dos planos de R^n son iguales, si y solo si, los dos planos son paralelos y tienen al menos un punto común.□

Es natural, que nos preguntemos ahora qué significa que una recta y un plano sean paralelos y qué significa que sean ortogonales, veamos la definición y revisémosla en un par de ejemplos.

Definición 21 [*Recta y plano paralelos*]. Sean $\mathcal{L}$ una recta con vector director $\mathbf{d} \in R^n$ y $\mathcal{P}$ un plano con vectores directores $\mathbf{c}_1$ y $\mathbf{d}_1 \in R^n$. Diremos que la recta $\mathcal{L}$ es paralela al plano $\mathcal{P}$, si y solo si, el vector $\mathbf{d}$ es combinación lineal de $\mathbf{c}_1$ y $\mathbf{d}_1$.



Ejemplo 34. Determinemos si la recta $\mathcal{L}$, cuya ecuación vectorial es $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + t \begin{pmatrix} 2 \\ -7 \\ -2 \end{pmatrix}$, $t \in R$, es paralela al plano $\mathcal{P}$, cuya ecuación vectorial es $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} + t \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} + s \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$, $t, \ s \in R$.
Un vector director de la recta $\mathcal{L}$ es $\mathbf{d} = \begin{pmatrix} 2 \\ -7 \\ -2 \end{pmatrix}$ y dos vectores directores del plano $\mathcal{P}$ son $\mathbf{c}_1 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$ y $\mathbf{d}_1 = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$. Veamos si $\mathbf{d}$ es combinación lineal de $\mathbf{c}_1$ y $\mathbf{d}_1$; es decir, si el sistema, cuya matriz aumentada es

$$\left(\begin{array}{rr|r} 0 & 2 & 2 \\ -2 & 0 & -7 \\ 1 & -3 & -2 \end{array}\right),$$

es consistente. Al escalonar la matriz, obtenemos

$$\left(\begin{array}{rr|r} 1 & -3 & -2 \\ 0 & -6 & -11 \\ 0 & 0 & -5/3 \end{array}\right),$$

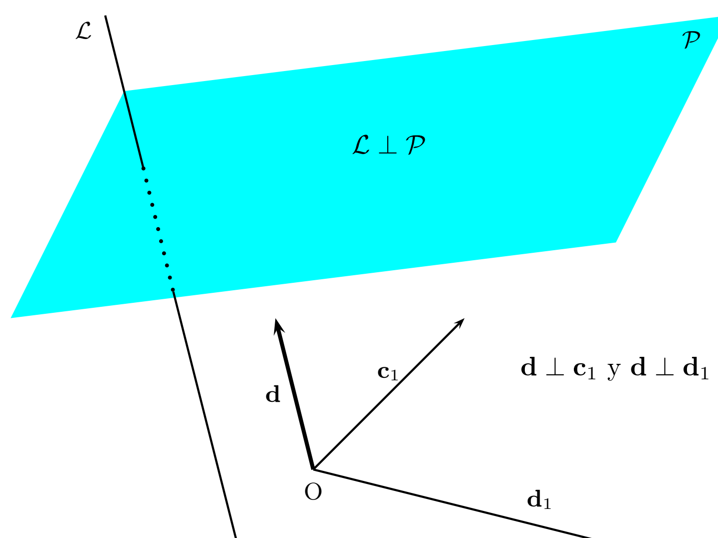
lo cual nos muestra que el sistema no tiene solución; es decir, que el vector director de la recta $\mathcal{L}$ no es combinación lineal de los vectores directores del plano $\mathcal{P}$. Por lo tanto, la recta $\mathcal{L}$ y el plano $\mathcal{P}$ no son paralelos. □

Una aplicación inmediata de esta definición es la caracterización de la inclusión de una recta en un plano de R^n, como lo planteamos en el siguiente teorema.

Teorema 15 [*Inclusión de una recta en un plano*].
Una recta está totalmente incluida en un plano de R^n, si y solo si, la recta es paralela al plano y la recta y el plano tienen al menos un punto en común.

Demostración:

Definición 22 [*Recta y plano ortogonales*]. Sean $\mathcal{L}$ una recta con vector director $\mathbf{d} \in R^n$ y $\mathcal{P}$ un plano con vectores directores $\mathbf{c}_1$ y $\mathbf{d}_1 \in R^n$. Diremos que la recta $\mathcal{L}$ es ortogonal al plano $\mathcal{P}$, si y solo si, el vector $\mathbf{d}$ es ortogonal tanto a $\mathbf{c}_1$, como a $\mathbf{d}_1$.



Ejemplo 35. Determinemos si la recta $\mathcal{L}$ que contiene a los puntos $P = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ y $Q = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}$ es ortogonal al plano que contiene al punto $M = \begin{pmatrix} 5 \\ -2 \\ 3 \end{pmatrix}$ y tiene vectores directores $\mathbf{c}_1 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$ y $\mathbf{d}_1 = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$.

Un vector director de la recta $\mathcal{L}$ es $\mathbf{d} = \overline{PQ} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$. Veamos si éste es ortogonal a los vectores $\mathbf{c}_1$ y $\mathbf{d}_1$. Para tal efecto, calculemos

$$\mathbf{d} \cdot \mathbf{c}_1 = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} = 0 - 2 + 2 = 0 \quad \text{y} \quad \mathbf{d} \cdot \mathbf{d}_1 = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix} = 6 + 0 - 6 = 0.$$

Así que la recta $\mathcal{L}$ y el plano $\mathcal{P}$ son ortogonales. □

Ahora, así como podemos estar interesados en todos los puntos que, a partir de un punto fijo, determinan vectores paralelos a un vector dado (los que forman una recta), podemos estar interesados en los puntos que, a partir de un punto fijo, determinan vectores ortogonales a un vector dado. A este conjunto de puntos lo llamaremos *hiperplano*.

Definición 23 [*Hiperplano*]. Dados un punto P y un vector no nulo $\mathbf{n}$, diremos que el conjunto formado por P y todos los puntos X que determinan vectores $\overline{PX}$ ortogonales a $\mathbf{n}$ es el *hiperplano* que contiene al punto P y es ortogonal al vector $\mathbf{n}$. Al vector $\mathbf{n}$ lo llamamos *vector normal* del hiperplano .

Teniendo en cuenta la definición de vectores ortogonales, tenemos que dado un punto P, los vectores $\overline{PX} = \overline{OX} - \overline{OP}$ ortogonales a $\mathbf{n}$ son los que hacen que $\overline{PX} \cdot \mathbf{n} = 0$. Así que, si llamamos $\mathbf{x}$ al vector $\overline{OX}$ y $\mathbf{p}$ al vector $\overline{OP}$, la ecuación

$$(\mathbf{x} - \mathbf{p}) \cdot \mathbf{n} = 0 \quad \text{equivalentemente} \quad \mathbf{x} \cdot \mathbf{n} = \mathbf{p} \cdot \mathbf{n} \tag{2.6}$$

es la ecuación del hiperplano que pasa por P y es ortogonal a $\mathbf{n}$. A esta ecuación la llamamos *ecuación normal del hiperplano*. Al expresar el primer producto escalar de (2.6), en término de las componentes de los vectores $\mathbf{n} = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$ y $\mathbf{x} - \mathbf{p} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, obtenemos

$$l_1(x_1 - a_1) + l_2(x_2 - a_2) + \ldots + l_n(x_n - a_n) = 0$$

ó equivalentemente,

$$l_1x_1 + l_2x_2 + \ldots + l_nx_n = d \qquad \text{con} \quad d = l_1a_1 + l_2a_2 + \ldots + l_na_n = \mathbf{n} \cdot \mathbf{p}.$$

A esta ecuación la llamamos *ecuación general del hiperplano* que pasa por P y es ortogonal a $\mathbf{n}$.

Ejemplo 36. Hallemos una ecuación del hiperplano que pasa por el punto $P = \begin{pmatrix} 2 \\ -3 \\ 5 \\ 1 \end{pmatrix}$ y es ortogonal al Eje X.

Un vector que tiene la dirección del Eje X es $\mathbf{e}_1$. Así que una ecuación para este plano es

$$1(x-2) + 0(y+3) + 0(z-5) + 0(w-1) = 0$$

o equivalentemente,

$$x - 2 = 0 \qquad \text{ó} \qquad x = 2.$$

□

Como sucedió con las rectas y sus vectores directores, el paralelismo y la ortogonalidad de dos hiperplanos depende del paralelismo y la ortogonalidad de sus vectores normales, respectivamente, como lo expresan las siguientes definiciones.

Definición 24 [*Hiperplanos paralelos*]. Sean $\mathcal{H}_1$ y $\mathcal{H}_2$ dos hiperplanos con vectores normales $\mathbf{n}_1$ y $\mathbf{n}_2$, respectivamente. Diremos que los hiperplanos $\mathcal{H}_1$ y $\mathcal{H}_2$ son paralelos, si y solo si, los vectores $\mathbf{n}_1$ y $\mathbf{n}_2$ son paralelos.

Ejemplo 37. Encontremos una ecuación del hiperplano $\mathcal{H}_1$ de R^4 que pasa por el origen y es paralelo al hiperplano $\mathcal{H}_2$ definido por $3x_1 - 2x_3 + x_4 = 5$.

Si $\mathcal{H}_1$ es paralelo a $\mathcal{H}_2$, podemos tomar como vector normal de $\mathcal{H}_1$ el mismo vector normal de $\mathcal{H}_2$, el cual tendrá como componentes los coeficientes respectivos de las variables en la ecuación general del hiperplano

$\mathcal{H}_2$; es decir, $\mathbf{n}_2 = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 1 \end{pmatrix}$. Como un punto de $\mathcal{H}_1$ es el origen, una ecuación del hiperplano es $(\mathbf{x}-\mathbf{0})\cdot\mathbf{n}_2 = 0$; es decir,

$$3x_1 - 2x_3 + x_4 = 0.$$

Existe otro hiperplano $\mathcal{H}_1$ que cumpla con las mismas condiciones? □

Definición 25 [*Hiperplanos ortogonales*]. Sean $\mathcal{H}_1$ y $\mathcal{H}_2$ dos hiperplanos con vectores normales $\mathbf{n}_1$ y $\mathbf{n}_2$, respectivamente. Diremos que los hiperplanos $\mathcal{H}_1$ y $\mathcal{H}_2$ son ortogonales, si y solo si, los vectores $\mathbf{n}_1$ y $\mathbf{n}_2$ son ortogonales.

Ejemplo 38. Encontremos una ecuación de un hiperplano $\mathcal{H}_1$ de R^5 que pasa por el origen y que sea ortogonal al hiperplano $\mathcal{H}_2$ definido por $3x_2 - x_3 - 2x_4 = 2$.

Si $\mathcal{H}_1$ es ortogonal a $\mathcal{H}_2$, el vector normal de $\mathcal{H}_1$, $\mathbf{n}_1$, debe ser ortogonal al vector normal de $\mathcal{H}_2$, el cual tendrá como componentes los coeficientes respectivos de las variables en la ecuación general de $\mathcal{H}_2$; es decir, $\mathbf{n}_2 = \begin{pmatrix} 0 \\ 3 \\ -1 \\ -2 \\ 0 \end{pmatrix}$; esto quiere decir que $\mathbf{n}_1$ debe ser tal que $\mathbf{n}_1 \cdot \mathbf{n}_2 = 0$; por lo tanto, podemos tomar $\mathbf{n}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix}$. Como un punto de $\mathcal{H}_1$ es el origen, su ecuación es $(\mathbf{x} - \mathbf{0}) \cdot \mathbf{n}_1 = 0$; es decir,

$$x_1 + 2x_5 = 0.$$

Existe otro hiperplano $\mathcal{H}_1$ que cumpla con las mismas condiciones?

□

Como casos especiales, en R, un hiperplano es un punto; en R^2, un hiperplano es una recta y, en R^3, los hiperplanos coinciden con los planos. Dejamos como ejercicio para el lector la demostración de los casos en R y en R^2. Para demostrar la afirmación en el caso de R^3, requerimos el cálculo de un vector ortogonal a otros dos, razón por la cual introducimos una operación cuya definición está **restringida** a R^3 y que llamamos *producto vectorial*.

Definición 26 [*Producto vectorial*].[9] Dados dos vectores $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ de R^3, definimos $\mathbf{u} \times \mathbf{v}$, el *producto vectorial* de $\mathbf{u}$ y $\mathbf{v}$, como el vector

$$\mathbf{u} \times \mathbf{v} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ -(u_1v_3 - u_3v_1) \\ u_1v_2 - u_2v_1 \end{pmatrix}.$$ [10]

[9] En la literatura, otro nombre que recibe esta operación es *producto cruz.*

[10] En el Capítulo 3, Sección 8, donde presentaremos la definición de determinantes, veremos que este producto lo podemos expresar en término de determinantes.

Ejemplo 39. Demostremos que $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ y $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$.[11]

$$\mathbf{e}_1 \times \mathbf{e}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 0 - 0 \cdot 1 \\ -(1 \cdot 0 - 0 \cdot 0) \\ 1 \cdot 1 - 0 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{e}_3$$

y de manera similar, tenemos

$$\mathbf{e}_2 \times \mathbf{e}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 - 0 \cdot 0 \\ -(0 \cdot 1 - 0 \cdot 0) \\ 0 \cdot 0 - 1 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \mathbf{e}_1$$

y

$$\mathbf{e}_3 \times \mathbf{e}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 0 - 1 \cdot 0 \\ -(0 \cdot 0 - 1 \cdot 1) \\ 0 \cdot 0 - 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \mathbf{e}_2$$

□

En el siguiente teorema, consignamos las principales propiedades algebraicas de este producto.

Teorema 16 [*Propiedades del producto vectorial*].
Si $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ son vectores de R^3 y λ es un escalar, entonces:

1. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$. *Ley anticonmutativa*
2. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$. *Ley distributiva para la suma por derecha*
3. $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$. *Ley distributiva para la suma por izquierda*
4. $\lambda(\mathbf{u} \times \mathbf{v}) = (\lambda\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (\lambda\mathbf{v})$.
5. $\mathbf{u} \times \mathbf{0} = \mathbf{0} \times \mathbf{u} = \mathbf{0}$.
6. $\mathbf{u} \times \mathbf{u} = \mathbf{0}$.
7. $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$.
8. $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{v} = 0$.
9. $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$.

Demostración: Demostremos las Propiedades 1 y 8. El resto de las demostraciones las dejamos como ejercicio para el lector. Para demostrar la Propiedad 1, tenemos que

$$\mathbf{u} \times \mathbf{v} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ -(u_1v_3 - u_3v_1) \\ u_1v_2 - u_2v_1 \end{pmatrix} = - \begin{pmatrix} -(u_2v_3 - u_3v_2) \\ (u_1v_3 - u_3v_1) \\ -(u_1v_2 - u_2v_1) \end{pmatrix} = - \begin{pmatrix} v_2u_3 - v_3u_2 \\ -(v_1u_3 - v_3u_1) \\ v_1u_2 - v_2u_1 \end{pmatrix} = -\mathbf{v} \times \mathbf{u}.$$

Para la demostración de la Propiedad 8, tenemos que

$$\begin{aligned} (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} &= \begin{pmatrix} u_2v_3 - u_3v_2 \\ -(u_1v_3 - u_3v_1) \\ u_1v_2 - u_2v_1 \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \\ &= (u_2v_3 - u_3v_2)u_1 - (u_1v_3 - u_3v_1)u_2 + (u_1v_2 - u_2v_1)u_3 \\ &= u_2v_3u_1 - u_3v_2u_1 - u_1v_3u_2 + u_3v_1u_2 + u_1v_2u_3 - u_2v_1u_3 \\ &= 0 \end{aligned}$$

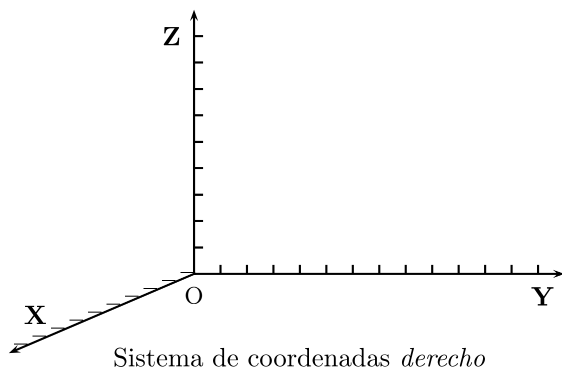
[11] Como anotamos antes, en la literatura, los vectores canónicos de R^3, $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ también son denotados como $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ y $\hat{\mathbf{k}}$, respectivamente. Asi, los resultados de este ejemplo se escribirían como $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}$, $\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}$ y $\hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}}$, respectivamente.

y de igual manera, tenemos que

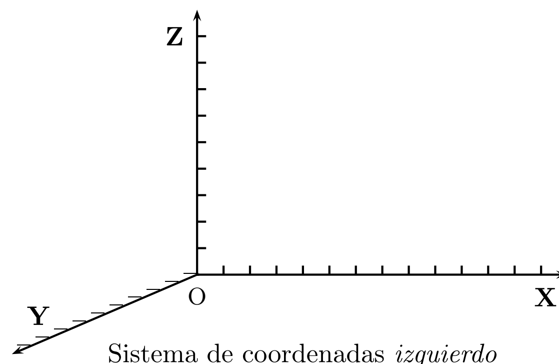
$$\begin{aligned}(\mathbf{u}\times\mathbf{v})\cdot\mathbf{v} &= \begin{pmatrix} u_2v_3-u_3v_2 \\ -(u_1v_3-u_3v_1) \\ u_1v_2-u_2v_1\end{pmatrix}\cdot\begin{pmatrix} v_1\\ v_2\\ v_3\end{pmatrix}\\ &= (u_2v_3-u_3v_2)v_1-(u_1v_3-u_3v_1)v_2+(u_1v_2-u_2v_1)v_3\\ &= u_2v_3v_1-u_3v_2v_1-u_1v_3v_2+u_3v_1v_2+u_1v_2v_3-u_2v_1v_3\\ &= 0\end{aligned}$$

□

Es importante anotar que la Propiedad 8 nos dice que el vector $\mathbf{u}\times\mathbf{v}$ siempre es ortogonal tanto a $\mathbf{u}$, como a $\mathbf{v}$, así que la dirección del vector $\mathbf{u}\times\mathbf{v}$ queda determinada por los vectores $\mathbf{u}$ y $\mathbf{v}$. El sentido de $\mathbf{u}\times\mathbf{v}$ depende del tipo de sistema de coordenadas: *derecho ó izquierdo*[12].



Sistema de coordenadas *derecho*



Sistema de coordenadas *izquierdo*

En un sistema de coordenadas derecho, el sentido de $\mathbf{u}\times\mathbf{v}$ está dado por el dedo pulgar de la mano derecha cuando ésta se coloca de tal forma que el resto de dedos apunten en el sentido de $\mathbf{u}$ y cierren hacia $\mathbf{v}$. Para terminar de caracterizar el vector $\mathbf{u}\times\mathbf{v}$, nos falta saber como calcular su norma, lo cual consignamos en el siguiente teorema.

Teorema 17 [*Norma del producto vectorial*].
Dados dos vectores arbitrarios $\mathbf{u}$ y $\mathbf{v}$ de R^3, si θ es el ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$, entonces tenemos las siguientes igualdades.

1. $\|\mathbf{u}\times\mathbf{v}\|^2=\|\mathbf{u}\|^2\|\mathbf{v}\|^2-(\mathbf{u}\cdot\mathbf{v})^2$. [*Identidad de Lagrange*]
2. $\|\mathbf{u}\times\mathbf{v}\|=\|\mathbf{u}\|\|\mathbf{v}\|\operatorname{sen}\theta$. [*Norma del producto vectorial*]

Demostración:

1. Usando las Propiedades 7 y 9 del teorema anterior y las propiedades del producto escalar, tenemos

$$\begin{aligned}\|\mathbf{u}\times\mathbf{v}\|^2 &= (\mathbf{u}\times\mathbf{v})\cdot(\mathbf{u}\times\mathbf{v})\\ &= \mathbf{u}\cdot[\mathbf{v}\times(\mathbf{u}\times\mathbf{v})] \quad \text{Propiedad 9, Teorema 16 con } \mathbf{w}=\mathbf{u}\times\mathbf{v}\\ &= \mathbf{u}\cdot[(\mathbf{v}\cdot\mathbf{v})\mathbf{u}-(\mathbf{v}\cdot\mathbf{u})\mathbf{v}] \quad \text{Propiedad 7, Teorema 16}\\ &= (\mathbf{v}\cdot\mathbf{v})(\mathbf{u}\cdot\mathbf{u})-(\mathbf{v}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{v})\\ &= \|\mathbf{v}\|^2\|\mathbf{u}\|^2-(\mathbf{u}\cdot\mathbf{v})^2\end{aligned}$$

[12] Un *sistema de coordenadas derecho* tiene definido los semiejes positivos de tal forma que, empezando con los dedos de la mano derecha (excepto el pulgar) en la dirección del semieje positivo de **X**, al cerrarlos hacia el semieje positivo de **Y**, el pulgar apunta en la dirección del semieje positivo de **Z**. Si ocurre lo mismo, pero usando la mano izquierda, tenemos un *sistema de coordenadas izquierdo*

2. Usando el anterior resultado, las propiedades del producto escalar y la identidad trigonométrica básica ($\cos^2\theta + \operatorname{sen}^2\theta = 1$), tenemos

$$\begin{aligned}
\|\mathbf{u}\times\mathbf{v}\|^2 &= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u}\cdot\mathbf{v})^2. \\
&= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \|\mathbf{u}\|^2\|\mathbf{v}\|^2\cos^2\theta \\
&= \|\mathbf{u}\|^2\|\mathbf{v}\|^2(1-\cos^2\theta) \\
&= \|\mathbf{u}\|^2\|\mathbf{v}\|^2\operatorname{sen}^2\theta
\end{aligned}$$

Por tanto, $\|\mathbf{u}\times\mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\|\operatorname{sen}\theta$. □

De este último resultado, obtenemos una nueva caracterización de vectores paralelos, la cual consignamos en el siguiente corolario. Sin embargo, notemos que, para determinar paralelismo entre vectores, es más práctico utilizar el criterio basado en el producto por escalar.

Corolario 17.1
Dos vectores no nulos de R^3 son paralelos, si y solo si, $\mathbf{u}\times\mathbf{v} = \mathbf{0}$.

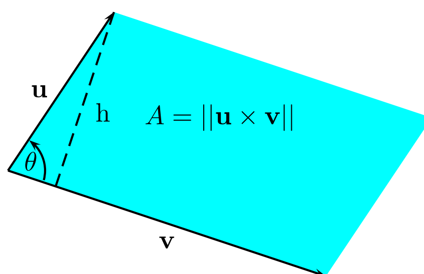
Demostración: Tenemos que $\mathbf{u}\times\mathbf{v} = \mathbf{0}$, si y solo si, $\|\mathbf{u}\times\mathbf{v}\| = 0$, pero $\|\mathbf{u}\times\mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\|\operatorname{sen}\theta$ y puesto que $\mathbf{u}$ y $\mathbf{v}$ son vectores no nulos, tenemos que $\operatorname{sen}\theta = 0$, por lo tanto $\theta = 0$ ó $\theta = \pi$, lo que implica que los vectores $\mathbf{u}$ y $\mathbf{v}$ son paralelos. De otro lado, si $\mathbf{u}$ es paralelo a $\mathbf{v}$, $\mathbf{v} = \lambda\mathbf{u}$, por las Propiedades 4 y 6 del Teorema 16,

$$\mathbf{u}\times\mathbf{v} = \mathbf{u}\times(\lambda\mathbf{u}) = \lambda(\mathbf{u}\times\mathbf{u}) = \lambda\mathbf{0} = \mathbf{0}.$$

□

Otra consecuencia del anterior teorema es la interpretación geométrica de su segundo resultado, la cual consignamos en el siguiente corolario.

Corolario 17.2 [*Área de un paralelogramo*].
El área del paralelogramo cuyos lados no paralelos están dados por los vectores $\mathbf{u}$ y $\mathbf{v}$ de R^3 está dada por la magnitud del producto vectorial de ellos, es decir, por $\|\mathbf{u}\times\mathbf{v}\|$.



Demostración: Consideremos el paralelogramo cuyos lados no paralelos son los vectores (libres) $\mathbf{u}$ y $\mathbf{v}$, y θ el ángulo entre ellos, como muestra la figura anterior.

Si observamos que h, la altura del paralelogramo, está dada por $h = \|\mathbf{u}\|\operatorname{sen}\theta$ y recordamos que A, el área del paralelogramo, es base por altura, tenemos

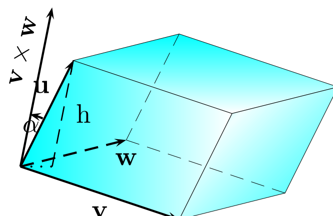
$$A = \|\mathbf{v}\|h = \|\mathbf{v}\|\|\mathbf{u}\|\operatorname{sen}\theta = \|\mathbf{u}\times\mathbf{v}\|,$$

según el Resultado 2 del Teorema 17. □

A su vez, este resultado nos permite calcular el volumen de un paralelepípedo en término de los vectores que definen sus tres aristas no paralelas, como lo muestra el siguiente corolario.

Corolario 17.3 [*Volumen de un paralelepípedo*].
El volumen del paralelepípedo cuyas aristas no paralelas están dadas por los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ de R^3 está dado por el valor absoluto del *producto mixto* de ellos, es decir, por $|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|$.

Demostración: Consideremos el paralelepípedo cuyas aristas no paralelas son los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ y sea α el ángulo entre $\mathbf{u}$ y $\mathbf{v}\times\mathbf{w}$, como muestra la figura (recordemos que el vector $\mathbf{v}\times\mathbf{w}$ es ortogonal a la base del paralelepípedo definida por $\mathbf{v}$ y $\mathbf{w}$)



Observemos que h, la altura del paralelepípedo, está dada por $h = \|\mathbf{u}\|\,|\cos\alpha|$ (h es paralela a $\mathbf{v}\times\mathbf{w}$) y recordemos que V, el volumen del paralelepípedo, es el área de la base por la altura. Por el resultado anterior, el área de la base es $\|\mathbf{v}\times\mathbf{w}\|$, por lo tanto tenemos

$$V = \|\mathbf{v}\times\mathbf{w}\|h = \|\mathbf{v}\times\mathbf{w}\|\|\mathbf{u}\||\cos\alpha| = |\|\mathbf{v}\times\mathbf{w}\|\,\|\mathbf{u}\|\cos\alpha| = |\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|.$$

□

Una consecuencia inmediata de este resultado es una caracterización de tres vectores linealmente independientes en R^3. Lo anterior es equivalente a determinar cuándo existe un plano que contenga a tres vectores dados (Ejercicio). En este caso, diremos que los tres vectores son *coplanares*.

Corolario 17.4 [*Vectores coplanares*].
Tres vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}\in R^3$ son coplanares, si y solo si, $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = 0$.

Demostración: El resultado lo obtenemos del hecho que tres vectores de R^3 son coplanares, si y solo si, el volumen del paralelepípedo formado por los tres vectores no tiene altura (o ésta es cero), por lo tanto, su volumen será cero. La ecuación la obtenemos del Corolario 17.3. □

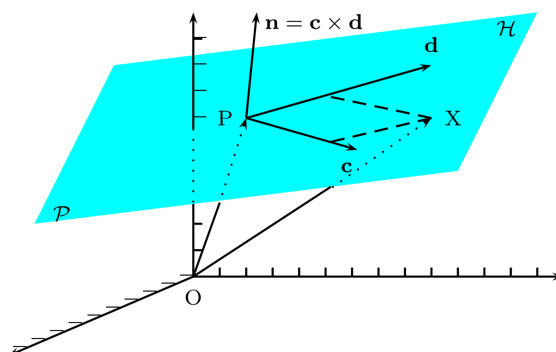
Finalmente, la Propiedad 8 del Teorema 16 y el anterior resultado nos permiten encontrar la ecuación normal de un plano a partir de su ecuación vectorial, al tiempo que nos muestra que un hiperplano en R^3 es un plano, como fue el propósito de introducir el producto vectorial, lo cual lo resumimos en el siguiente teorema.

Teorema 18 [*Ecuación normal del plano en* R^3].
El plano $\mathcal{P}$ de R^3 que contiene al punto P y tienen vectores directores $\mathbf{c}$ y $\mathbf{d}$, y el hiperplano $\mathcal{H}$ de R^3 que contiene el punto P y es ortogonal a $\mathbf{n} = \mathbf{c}\times\mathbf{d}$ son iguales.

Demostración: Sea $X\in\mathcal{P}$. Por la definición de plano, $\overline{PX}$ es una combinación lineal de $\mathbf{c}$ y $\mathbf{d}$, es decir, existen escalares α y β tales que $\overline{PX} = \alpha\mathbf{c}+\beta\mathbf{d}$. Por la Propiedad 8 del Teorema 16,

$$\overline{PX}\cdot\mathbf{n} = (\alpha\mathbf{c}+\beta\mathbf{d})\cdot(\mathbf{c}\times\mathbf{d}) = \alpha\mathbf{c}\cdot(\mathbf{c}\times\mathbf{d}) + \beta\mathbf{d}\cdot(\mathbf{c}\times\mathbf{d}) = 0,$$

por lo tanto, $\overline{PX}$ es ortogonal a $\mathbf{n}$, de donde, por la definición de hiperplano, concluimos que $X\in\mathcal{H}$.



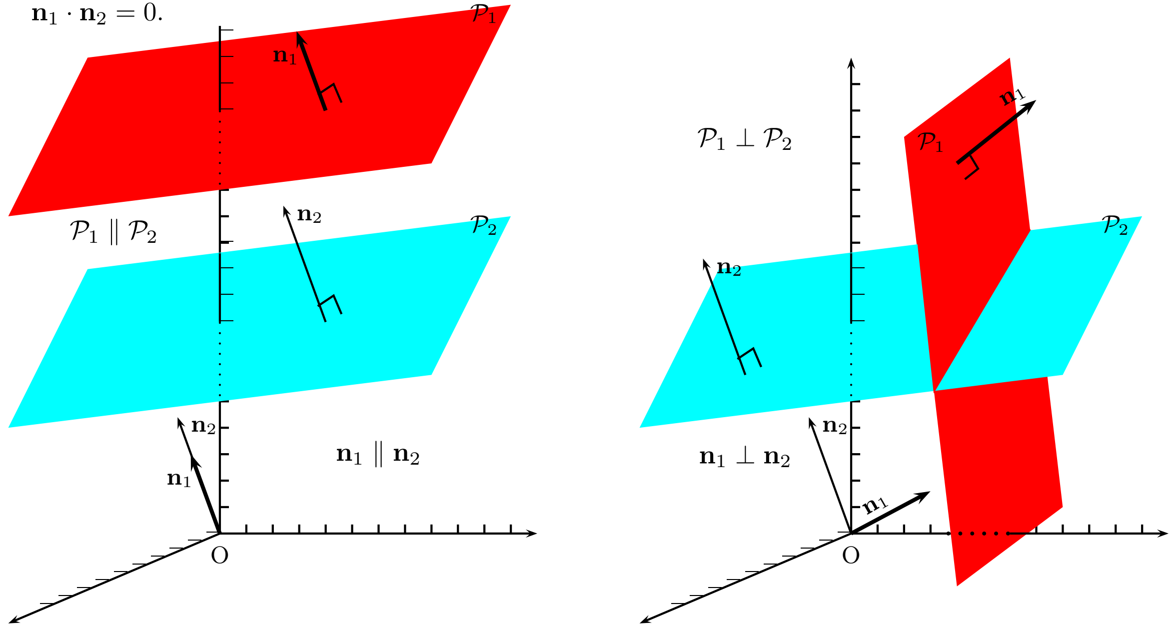
Ahora, sea $X \in \mathcal{H}$. Por la definición de hiperplano, $\overline{PX}$ es ortogonal a $\mathbf{n} = \mathbf{c} \times \mathbf{d}$, por lo tanto $\overline{PX} \cdot (\mathbf{c} \times \mathbf{d}) = 0$, lo cual , por el Corolario 17.4, significa que $\overline{PX}$, $\mathbf{c}$ y $\mathbf{d}$ son coplanares. Como $\mathbf{c}$ y $\mathbf{d}$ son vectores diferentes de cero y no paralelos (Por qué?), existen escalares α y β tales que $\overline{PX} = \alpha\mathbf{c} + \beta\mathbf{d}$. Por la definición de plano, concluimos que $X \in \mathcal{P}$. □

Este resultado no solo nos permite otra forma de caracterizar los puntos que conforman un plano (Ecuación normal del plano), sino que también enriquece tanto el análisis como el cálculo en la Geometría Euclidiana del espacio (o en R^3) como lo planteamos en los siguientes teoremas sobre paralelismo y ortogonalidad entre planos, y entre rectas y planos en el espacio, donde algunas demostraciones las dejamos como ejercicio para el lector.

Teorema 19 [*Paralelismo y ortogonalidad de planos en* R^3].
Sean $\mathcal{P}_1$ y $\mathcal{P}_2$ dos planos en R^3, con vectores normales $\mathbf{n}_1$ y $\mathbf{n}_2$, respectivamente.

1. Los planos $\mathcal{P}_1$ y $\mathcal{P}_2$ son paralelos, si y solo si, los vectores $\mathbf{n}_1$ y $\mathbf{n}_2$ lo son; es decir, si y solo si, $\mathbf{n}_1$ es múltiplo por escalar de $\mathbf{n}_2$.

2. Los planos $\mathcal{P}_1$ y $\mathcal{P}_2$ son ortogonales, si y solo si, los vectores $\mathbf{n}_1$ y $\mathbf{n}_2$ lo son; es decir, si y solo si, $\mathbf{n}_1 \cdot \mathbf{n}_2 = 0$.



Demostración: La demostración de estas dos proposiciones se siguen del resultado del Teorema 18 (en R^3, un plano es un hiperplano) y las Definiciones 24 y 25, sobre paralelismo y ortogonalidad de hiperplanos. □

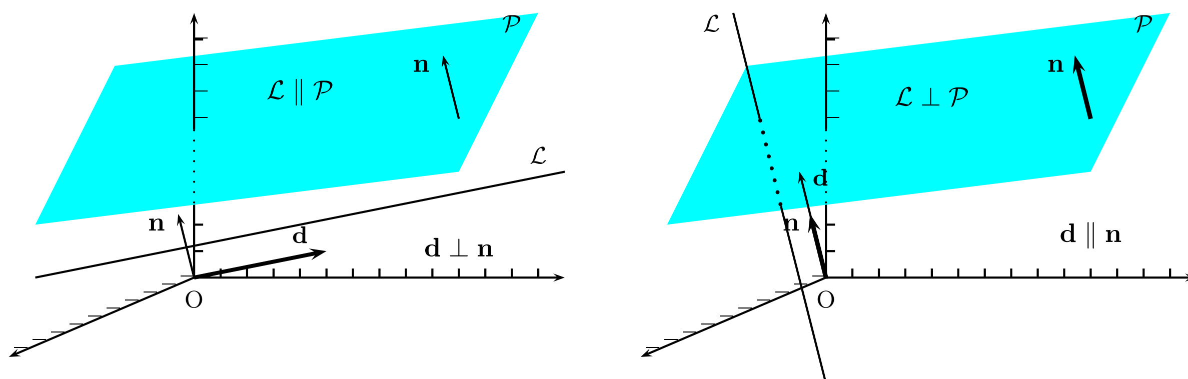
Ejemplo 40. Determinemos si el plano $\mathcal{P}_1$ del Ejemplo 33 es paralelo o es ortogonal al plano $\mathcal{P}_2$ definido por $3x - z = 4$.

Por el Teorema 18, un vector normal de $\mathcal{P}_1$ es $\mathbf{n}_1 = \mathbf{c}_1 \times \mathbf{d}_1 = \begin{pmatrix} 2 \\ -7 \\ -2 \end{pmatrix} \times \begin{pmatrix} 4 \\ -5 \\ -6 \end{pmatrix} = \begin{pmatrix} 32 \\ 4 \\ 18 \end{pmatrix}$ y, por la definición de hiperplano, un vector normal de $\mathcal{P}_2$ es $\mathbf{n}_2 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}$. Como no existe un escalar λ tal que $\mathbf{n}_1 = \lambda\mathbf{n}_2$, estos planos no son paralelos, y como $\mathbf{n}_1 \cdot \mathbf{n}_2 = 78 \neq 0$, tampoco son ortogonales. □

Teorema 20 [*Paralelismo y ortogonalidad de rectas y planos en* R^3].
Sean $\mathcal{L}$ una recta con vector director $\mathbf{d} \in R^3$ y $\mathcal{P}$ un plano con vector normal $\mathbf{n} \in R^3$.

1. La recta $\mathcal{L}$ es **paralela** al plano $\mathcal{P}$, si y solo si, el vector $\mathbf{d}$ es **ortogonal** al vector $\mathbf{n}$; es decir, si y solo si, $\mathbf{d} \cdot \mathbf{n} = 0$.

2. La recta $\mathcal{L}$ es **ortogonal** al plano $\mathcal{P}$, si y solo si, el vector **d** es **paralelo** al vector **n**; es decir, si y solo si, **d** es múltiplo por escalar de **n**.



Demostración:

1. Si la recta $\mathcal{L}$ y el plano $\mathcal{P}$ son paralelos, por la Definición 20, el vector **d** es combinación lineal de los vectores directores del plano $\mathcal{P}$; es decir, existen puntos $P, Q \in \mathcal{P}$ tales que $\mathbf{d} = \overline{PQ}$; finalmente, por la definición de hiperplano, **d** es ortogonal a **n**; es decir, $\mathbf{d} \cdot \mathbf{n} = 0$.

   Observemos que la anterior argumentación está basada en definiciones, por lo tanto, es también válida en el sentido contrario.

2. Ejercicio para el lector. □

Ejemplo 41. Determinemos si la recta $\mathcal{L}$ del Ejemplo 34 es paralela u ortogonal al plano $\mathcal{P}$ del Ejemplo 35.

Por el Teorema 18, un vector normal de $\mathcal{P}$ es $\mathbf{n} = \mathbf{c}_1 \times \mathbf{d}_1 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix}$. Como no existe un escalar λ tal que $\mathbf{d} = \begin{pmatrix} 2 \\ -7 \\ 2 \end{pmatrix} = \lambda\mathbf{n}$ (d es el vector director de la recta $\mathcal{L}$), la recta $\mathcal{L}$ no es ortogonal al plano $\mathcal{P}$, y como $\mathbf{d} \cdot \mathbf{n} = 6 \neq 0$, la recta $\mathcal{L}$ tampoco es paralela al plano $\mathcal{P}$. □

2.9. Ejercicios

1. Haga una lista de por lo menos 3 variables ó cantidades tales que

   a) se identifiquen como escalares.

   b) se identifiquen como vectores (coordenadas).

   c) se identifiquen como vectores libres.

   En los dos últimos casos, las variables pertenecen a R^n para algún valor de n. Indique el valor de n.

2. Represente geométricamente los vectores **c**=2**a**, **d**=**a**+**b** **e**=**c**-**b** **f**=3**b**+1.5**a** y **g**=3**a**-2**b** y diga cuales son **paralelos**, cuando

   a) $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ b) $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$

3. Para los vectores **a** y **b** del punto anterior, calcule

   a) $\frac{1}{2}\mathbf{a} - \frac{3}{2}\mathbf{a}$ b) $\mathbf{b} - 3\mathbf{a} - \mathbf{b}$ c) $\frac{3}{4}\mathbf{a} + \frac{3}{4}\mathbf{b}$

   Cuál de ellos es paralelo a **a**? cuál a **b**?

4. Trace dos vectores arbitrarios $\mathbf{u}, \mathbf{v}$ y construya gráficamente los vectores $2\mathbf{u}$, $-\dfrac{2}{3}\mathbf{u}$, $\mathbf{u} + 2\mathbf{v}$, $3\mathbf{u} - \mathbf{v}$. Cual de ellos es paralelo a **u**? cual a **v**?

5. En los torneos mixtos de fútbol, los equipos están conformados por 4 mujeres y 7 hombres, y en los de basketball, por 2 mujeres y 3 hombres, con la condición que cada equipo tenga un primíparo de cada sexo. Sean **f** y **b** los vectores que representan la composición por sexos de un equipo de fútbol y uno de basketball, respectivamente. Represente algebraica y geométricamente:

   a) La participación de los primíparos.

   b) La composición de los no primíparos en cada deporte.

   c) La reunión de 2 equipos de fútbol y 3 de basketball.

6. Un avión que vuela de sur a norte a una velocidad de 600 km/hora entra en una corriente de aire que va de este a oeste a una velocidad de 300 km/hora.

   a) Represente geométricamente estas dos cantidades y el *rumbo* final del avión.

   b) Gráficamente, Cual seria el *rumbo* final del avión si:

      1) La misma corriente va de oeste a este?
      2) La velocidad del avión hubiese sido la tercera parte?
      3) El avión vuela hacia el noreste?
      4) Las dos velocidades hubiesen sido el doble?
      5) La velocidad del avión hubiese sido la mitad y la de la corriente el doble?

7. Determine un vector **x** que satisfaga la ecuación

   a) $\begin{pmatrix} 1 \\ -4 \\ 0 \\ 2 \end{pmatrix} - 3\mathbf{x} = \begin{pmatrix} -2 \\ 0{,}5 \\ -1 \\ 5 \end{pmatrix}$ b) $\mathbf{x} + \dfrac{2}{3}\begin{pmatrix} 3 \\ -9 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 5 \\ -1 \end{pmatrix}$

8. Cuáles de las propiedades, clausurativa, conmutativa, asociativa, modulativa, opuestos y cancelativa de la suma entre números reales tienen una propiedad similar en la suma de vectores?. Cuáles de estas propiedades de la multiplicación entre números reales tienen una propiedad similar en la multiplicación por escalar?

9. Complete la demostración del Teorema 1.

10. Determine, para cada caso, los valores de a y b, si existen, que hacen válida la igualdad.

   a) $2\begin{pmatrix} -2 \\ 5 \\ -1 \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ b \end{pmatrix}$ b) $\begin{pmatrix} a-b \\ 0 \end{pmatrix} - \begin{pmatrix} a+3b \\ 2a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

11. Trace el vector $\overline{PQ}$ y calcule sus componentes.

   a) $P = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ y $Q = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ b) $P = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ y $Q = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$

12. Determine si el primer vector es combinación lineal de los otros.

a) $\begin{pmatrix} -9 \\ -4 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ -1 \end{pmatrix}$ b) $\begin{pmatrix} -2b \\ a+5b \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 5 \end{pmatrix}$

c) $\begin{pmatrix} 2a-2b \\ -a+6b \\ 5a-b \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$ d) $\begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \\ -3 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 1 \\ -1 \end{pmatrix}$

13. Determine si el vector $\mathbf{b}$ es combinación lineal de las columnas de la matriz dada.

a) $\mathbf{b} = \begin{pmatrix} -9 \\ -4 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 & 3 \\ 2 & 5 \\ 0 & -1 \end{pmatrix}$ b) $\mathbf{b} = \begin{pmatrix} -2b \\ a+5b \end{pmatrix}, \begin{pmatrix} -1 & -2 \\ 3 & 5 \end{pmatrix}$

14. Verifique que cualquier vector de R^3 es combinación lineal de $\begin{pmatrix} 3 \\ 3 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ y $\begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}$. Podemos afirmar que R^3 es generado por estos vectores?

15. Encuentre un vector de R^4 que no sea combinación lineal de $\begin{pmatrix} 3 \\ 0 \\ -3 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 1 \\ -2 \\ 3 \\ 0 \end{pmatrix}$. Podemos afirmar que R^4 es generado por estos vectores?

16. Determine para qué valores de λ, el vector $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ es combinación lineal de $\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\lambda \\ \lambda \end{pmatrix}$ y $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$.

17. Determine para qué valores de α, $Gen\left\{\begin{pmatrix} 0 \\ 1 \\ \alpha \end{pmatrix}, \begin{pmatrix} \alpha \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \alpha \\ 1 \end{pmatrix}\right\} = R^3$.

18. Dados los vectores $\mathbf{a} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$ y $\mathbf{c} = \begin{pmatrix} 0 \\ 12 \\ -24 \end{pmatrix}$.

a) Encuentre la combinación lineal $2\mathbf{a} - \mathbf{b} + \frac{1}{4}\mathbf{c}$.

b) Encuentre otra combinación lineal de los vectores dados.

c) Cuántas combinaciones lineales de los vectores dados existen?

d) Calcule los vectores $\mathbf{a} + \mathbf{b}$ y $\mathbf{a} + \mathbf{b} + \mathbf{c}$ y diga si son combinación lineal de los vectores dados.

e) Es el vector cero combinación lineal de los vectores dados?

f) Es el vector $\mathbf{u} = 2\mathbf{a} - 3\mathbf{b} + \frac{2}{3}\mathbf{c}$ combinación lineal de los vectores dados?

g) Es el vector $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ combinación lineal de los vectores dados?

h) Qué problema se plantea para contestar la pregunta anterior? Es necesario resolverlo?

i) El vector $\mathbf{a}$ pertenece a $Gen\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$?

j) El vector $\mathbf{a} + \mathbf{b}$ pertenece a $Gen\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$?

k) El vector $\mathbf{u}$ dado en f) pertenece a $Gen\{\mathbf{a},\mathbf{b},\mathbf{c}\}$?

l) El vector $\mathbf{v}$ dado en g) pertenece a $Gen\{\mathbf{a},\mathbf{b},\mathbf{c}\}$?

m) El vector $\mathbf{0}$ pertenece a $Gen\{\mathbf{a},\mathbf{b},\mathbf{c}\}$?

19. Reescriba cada una de las preguntas del punto anterior usando el producto $A\mathbf{x}$, indicando, en cada caso, cual es la matriz A y el vector $\mathbf{x}$.

20. Demuestre la segunda parte del Teorema 2.

21. Dados $\mathbf{u} = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 0 \\ 2 \\ -5 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$ y $H = \{\mathbf{u},\mathbf{v}\}$, determine cuales de las siguientes proposiciones son verdaderas.

a) El vector $\mathbf{v}$ está en H.

b) El vector $\mathbf{w}$ está en H.

c) El vector $\mathbf{v}$ está en $Gen\,\{H\}$.

d) El vector $\mathbf{w}$ está en $Gen\,\{H\}$.

e) El vector $\mathbf{2u\text{-}v}$ está en $Gen\,\{H\}$.

f) El vector $\mathbf{u}+\mathbf{3w}$ está en $Gen\,\{H\}$.

22. Explique por qué, si el conjunto M contiene un vector no nulo, $Gen\,\{M\}$ tiene infinitos vectores.

23. Dado un vector $\mathbf{u}$ de R^2, geométricamente, qué es $Gen\{\mathbf{u}\}$? $Gen\{\mathbf{u}, 2\mathbf{u}\}$?

24. Dados los vectores $\mathbf{u},\mathbf{v}$ de R^2, geométricamente, qué es $Gen\{\mathbf{u},\mathbf{v}\}$? qué es $Gen\{\mathbf{u},\mathbf{v},\mathbf{u}+\mathbf{v}\}$?

25. Dados los vectores $\mathbf{u},\mathbf{v}$ de R^3, geométricamente, qué es $Gen\{\mathbf{u},\mathbf{v}\}$? qué es $Gen\{\mathbf{u},\mathbf{v},\mathbf{u}+\mathbf{v}\}$?

26. Escriba un conjunto generador de $Gen\{\mathbf{u}, 2\mathbf{u}\}$. Existe otro conjunto generador con menos elementos?

27. Escriba un conjunto generador de $Gen\{\mathbf{u},\mathbf{v}\}$. Existe otro conjunto generador con menos elementos?

28. Escriba un conjunto generador de $Gen\{\mathbf{u},\mathbf{v},\mathbf{u}+\mathbf{v}\}$. Existe otro conjunto generador con menos elementos?

29. Verifique que $Gen\{\mathbf{u},\mathbf{v}\} = Gen\{\mathbf{u}+\mathbf{v},\mathbf{u}-\mathbf{v}\}$.

30. Dados los sistemas de ecuaciones lineales,

$$(i)\ \begin{aligned} 2x - y + 3z - w &= 0 \\ x + 3y - 5z &= 1 \end{aligned} \qquad (ii)\ \begin{aligned} x - 3z &= 1 \\ 3x + 2y &= -2 \\ 2y - z &= 0 \end{aligned} \qquad (iii)\ \begin{aligned} x_1 &= -2 \\ x_2 - 3x_3 &= 6 \\ x_2 - 3x_3 &= 6 \end{aligned}$$

a) Exprese los sistemas en forma de ecuaciones vectoriales

b) Exprese los sistemas como $A\mathbf{x} = \mathbf{b}$, indicando, en cada caso, cuál es la matriz A y cuales son los vectores $\mathbf{x}$ y $\mathbf{b}$.

31. Demuestre el Teorema 3 (AYUDA: Demuestre que cada uno de los puntos implica el siguiente y que el último implica el primero).

32. Sean A una matriz de m filas y n columnas y U una matriz escalonada equivalente a A. Si para cualquier vector $\mathbf{b}$ de R^m, el sistema de ecuaciones lineales, cuya matriz aumentada es $[A|\mathbf{b}]$, **tiene solución única**, determine cuales de las siguientes afirmaciones son verdaderas. Justifique su respuesta.

a) El vector $\mathbf{b}$ es combinación lineal de las columnas de A.

b) El vector **b** es combinación lineal de las columnas de U.

c) Cada fila de U tiene un pivote.

d) Cada columna de U tiene un pivote.

e) La matriz U tiene n pivotes.

f) La matriz U tiene m pivotes.

g) $m = n$.

h) Las columnas de A generan a R^m

33. Sean A una matriz de m filas y n columnas y U una matriz escalonada equivalente a A. Si para cualquier vector **b** de R^m, el sistema de ecuaciones lineales, cuya matriz aumentada es $[A|\mathbf{b}]$, **tiene infinitas soluciones**, determine cuales de las siguientes afirmaciones son verdaderas. Justifique su respuesta.

    a) El vector **b** es combinación lineal de las columnas de A.

    b) El vector **b** es combinación lineal de las columnas de U.

    c) Cada fila de U tiene un pivote.

    d) Cada columna de U tiene un pivote.

    e) La matriz U tiene n pivotes.

    f) La matriz U tiene m pivotes.

    g) $m < n$.

    h) Las columnas de A generan a R^m

34. Sean A una matriz de m filas y n columnas y U una matriz escalonada equivalente a A. Si para *un* vector **b** de R^m, el sistema de ecuaciones lineales, cuya matriz aumentada es $[A|\mathbf{b}]$, **es inconsistente**, determine cuales de las siguientes afirmaciones son verdaderas y responda a las preguntas formuladas. Justifique su respuesta.

    a) El vector **b** es combinación lineal de las columnas de A.

    b) El vector **b** es combinación lineal de las columnas de U.

    c) Cada fila de U tiene un pivote.

    d) El vector **b** puede ser **0**.

    e) El vector **b** puede ser un múltiplo de alguna de las columnas de A?

    f) El vector **b** puede ser la suma de las columnas de A?

    g) Las columnas de A generan a R^m.

    h) Qué puede decirse del número de pivotes de U?

35. Dados los vectores $\mathbf{a} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ -3 \end{pmatrix}$, $\mathbf{c} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ y la matriz $A = \begin{pmatrix} -1 & 3 & -5 \\ 2 & 5 & -1 \\ 0 & -1 & 1 \\ -2 & 0 & -41 \end{pmatrix}$

    a) Para que valor de n, el vector $\mathbf{z} = A\mathbf{b}$ pertenece a R^n?

    b) El vector **a** pertenece al espacio nulo de A? Al espacio columna de A?

    c) El vector cero pertenece al espacio nulo de A? Al espacio columna de A?

    d) El vector $\mathbf{v} = A\mathbf{b}$ pertenece al espacio nulo de A? Al espacio columna de A?

    e) El vector **c** pertenece al espacio columna de A? Al espacio nulo de A?

36. Determine si $Gen\left\{\begin{pmatrix}-1\\0\\3\end{pmatrix},\begin{pmatrix}2\\-5\\2\end{pmatrix}\right\}=Gen\left\{\begin{pmatrix}0\\-5\\8\end{pmatrix},\begin{pmatrix}1\\-5\\5\end{pmatrix}\right\}$.

37. Dadas las siguientes matrices,

i) $\begin{pmatrix}-3&3&-5\\2&5&-1\\1&-8&6\\5&2&4\end{pmatrix}$ ii) $\begin{pmatrix}2&5&-1\\-3&2&3\\5&3&-2\end{pmatrix}$ iii) $\begin{pmatrix}-2&5\\1&3\\5&3\\3&-2\end{pmatrix}$.

a) Determine si las columnas son vectores *l.i.*

b) Describa el problema que usted plantea para contestar la pregunta anterior.

c) Es necesario resolver el problema anterior?

d) Describa la pregunta a) en términos del espacio nulo de las matrices.

38. Para qué valores de r y t los siguientes conjuntos de vectores son *l.i.*?

a) $\left\{\begin{pmatrix}-r\\3\\t\end{pmatrix}\right\}$. b) $\left\{\begin{pmatrix}t-r\\2\end{pmatrix},\begin{pmatrix}-3\\1\end{pmatrix}\right\}$. c) $\left\{\begin{pmatrix}3r\\t\end{pmatrix},\begin{pmatrix}-r\\2r-t\end{pmatrix},\begin{pmatrix}r\\t\end{pmatrix}\right\}$.

39. Cuales de los siguientes conjuntos de vectores son *l.i.*? $\mathbf{u}\neq\mathbf{0}\neq\mathbf{v}$

a) $\{\mathbf{u},\mathbf{v},\mathbf{0}\}$. b) $\{\mathbf{u},\mathbf{v},2\mathbf{u}\}$. c) $\{\mathbf{u},\mathbf{v},\mathbf{u}-\mathbf{v}\}$. d) $\{\mathbf{u},\mathbf{v},\ \textit{tales que}\ \ \mathbf{u}\cdot\mathbf{v}=0\}$.

40. Complete la demostración del Teorema 6.

41. Dados los vectores $\mathbf{u},\mathbf{v}$ y $\mathbf{w}$ de R^3, señale las expresiones que están bien definidas.

a) $(\mathbf{u}\cdot\mathbf{v})\cdot\mathbf{w}$ b) $(\mathbf{u}\cdot\mathbf{v})\mathbf{w}$ c) $(\mathbf{u}\cdot\mathbf{v})(\mathbf{u}\cdot\mathbf{w})$ d) $(\mathbf{u}+\mathbf{v})\cdot\mathbf{w}$

e) $(\mathbf{u}\cdot\mathbf{v})+\mathbf{w}$ f) $(\mathbf{u}\cdot\mathbf{v})+3$ g) $(\mathbf{u}\cdot\mathbf{v})^{-2}$ h) $(\mathbf{u}\cdot\mathbf{v})/(\mathbf{u}\cdot\mathbf{w})$

42. Supongamos que la evaluación definitiva de un curso de Álgebra Lineal se determina por el promedio de parciales con un peso del 60 % , el promedio de quices con un peso del 30 % y una nota conceptual con un peso del 10 %. Si las calificaciones del estudiante Martin Pérez son 4.0 en el promedio de parciales, 4.5 en el promedio de quices y 4.8 en la nota conceptual, calcule la nota definitiva de Martin Pérez, usando el producto escalar.

43. Demuestre el Teorema 7.

44. Encuentre el ángulo que forman los siguientes vectores.

a) $\begin{pmatrix}-1\\0\\0\\1\end{pmatrix}$ y $\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$. b) $\begin{pmatrix}-2\\1\\\sqrt{3}\end{pmatrix}$ y $\begin{pmatrix}-2\\\sqrt{3}\\1\end{pmatrix}$. c) $\begin{pmatrix}-3\\3\end{pmatrix}$ y $\begin{pmatrix}\sqrt{2}\\\sqrt{2}\end{pmatrix}$.

45. Determine si los siguientes vectores son ortogonales

a) $\begin{pmatrix}-1\\2\\5\end{pmatrix}$ y $\begin{pmatrix}4\\-3\\2\end{pmatrix}$. b) $\begin{pmatrix}-7\\0\end{pmatrix}$ y $\begin{pmatrix}0\\-3\end{pmatrix}$. c) $\begin{pmatrix}4\\6\\2\\6\end{pmatrix}$ y $\begin{pmatrix}-1\\2\\5\\-3\end{pmatrix}$.

46. Encuentre, si existen, dos vectores ortogonales a los vectores dados

a) $\begin{pmatrix}-1\\5\\4\\-3\end{pmatrix}$ y $\begin{pmatrix}1\\0\\2\\-1\end{pmatrix}$. b) $\begin{pmatrix}-2\\3\end{pmatrix}$ y $\begin{pmatrix}0\\-5\end{pmatrix}$. c) $\begin{pmatrix}3\\6\\-1\end{pmatrix}$ y $\begin{pmatrix}-1\\2\\-3\end{pmatrix}$.

47. Dados los vectores

i) $\mathbf{u} = \begin{pmatrix} -5 \\ 2 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ y $\mathbf{w} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$. ii) $\mathbf{u} = \begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix}$ y $\mathbf{w} = \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix}$,

calcule (en el caso de vectores de R^2, de ser posible, ilustre sus respuestas en una gráfica)

a) $\mathbf{u} \cdot \mathbf{v}$, $\frac{1}{2}\mathbf{u} \cdot \frac{7}{2}\mathbf{v}$, $\mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}$ y $0{,}36\mathbf{v} - 0{,}36\mathbf{u}$.

b) la norma de: $\mathbf{u}$, $3\mathbf{u}$, $2\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$.

c) el ángulo entre $\mathbf{u}$ y $\mathbf{v}$, $\mathbf{u}$ y $2\mathbf{v}$, $\mathbf{v}$ y $-3\mathbf{v}$, $\mathbf{u}$ y $\mathbf{w}$, y $\mathbf{u}$ y $\mathbf{v} + \mathbf{w}$.

d) un vector unitario en la dirección y sentido de $\mathbf{u}$. Existe otro?

e) un vector paralelo a $\mathbf{v}$, con la mitad de su magnitud. Existe otro?

f) un vector en la dirección de $\mathbf{w}$ y sentido contrario a él. Existe otro?

g) un vector unitario ortogonal a $\mathbf{u}$. Existe otro?

h) un vector ortogonal a $\mathbf{u}$ y $\mathbf{v}$. Existe otro?.

i) un vector ortogonal a $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$. Existe otro?.

j) la proyección del vector $\mathbf{u}$ sobre el vector $\mathbf{v}$.

k) la proyección del vector $\mathbf{v}$ sobre el vector $\mathbf{u}$. Compare su respuesta con la del item anterior.

l) la proyección del vector $\mathbf{u}$ sobre el vector $2\mathbf{v}$. Compare su respuesta con la del item *j*).

m) la componente vectorial de $\mathbf{u}$ ortogonal a $2\mathbf{v}$.

n) el punto definido por un vector paralelo a $\mathbf{v}$ más cercano al punto $\mathbf{u}$. Compare su respuesta con la del item *j*).

48. Determine cuales de las siguientes afirmaciones son verdaderas.

a) Si $\mathbf{u} \cdot \mathbf{v} = 0$, entonces $\mathbf{u} = \mathbf{0}$ ó $\mathbf{v} = \mathbf{0}$.

b) Si $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$, entonces $\mathbf{v} = \mathbf{w}$.

c) Cualquier vector de R^n es *l.i.*

d) Cualquier par de vectores diferentes de R^n son *l.i.*

e) Cualquier tres vectores diferentes de R^3 son *l.i.*

f) Cualquier tres vectores diferentes de R^2 generan a R^2.

g) Cualquier par de vectores *l.i.* de R^2 generan a R^2.

h) Si el conjunto de vectores $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es *l.i.*, entonces el conjunto de vectores $\{\mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} + \mathbf{w}\}$ es *l.i.*

i) Si A es una matriz $m \times n$, cuyas columnas son vectores *l.i.*, entonces el sistema, cuya matriz aumentada asociada es $[A|\mathbf{b}]$, tiene solución para cualquier vector $\mathbf{b}$ de R^m.

j) Si A es una matriz $n \times n$, cuyas columnas son vectores *l.i.*, entonces el sistema, cuya matriz aumentada asociada es $[A|\mathbf{b}]$, tiene solución única para cualquier vector $\mathbf{b}$ de R^n.

k) Si el vector $\mathbf{u}$ es ortogonal a los vectores $\mathbf{v}$ y $\mathbf{w}$, entonces $\mathbf{u}$ es ortogonal a cualquier combinación lineal no nula de $\mathbf{v}$ y $\mathbf{w}$.

49. Dadas las siguientes ecuaciones, identifique las rectas, los planos y los hiperplanos

a) En R^2, $\left[\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 3 \\ -5 \end{pmatrix}\right] \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 0$

b) En R^4, $\begin{cases} x = 3 - t \\ y = t \\ z = 2 + 5t \\ w = 0 \end{cases}$, $t \in R$

c) En R^5, $\dfrac{x_1 - 3}{2} = \dfrac{x_2}{5} = x_3 = x_5 - 2, \quad x_4 = 0$

d) En R^4, $3x - 2y = 5$

e) En R^3, $\begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$

f) En R^4, $\begin{cases} x_1 = 2 - 3t \\ x_2 = -s \\ x_3 = 5 \\ x_4 = t + s \end{cases}$, $s, t \in R$

50. Encuentre dos puntos y dos vectores directores de la recta de R^4, $\mathcal{L}: \dfrac{x+5}{-3} = 1 + z = \dfrac{w}{2}, \quad y = 0.$

51. Encuentre una ecuación de una recta que contenga los puntos $P = \begin{pmatrix} 3 \\ -5 \\ 0 \\ 2 \end{pmatrix}$ y $Q = \begin{pmatrix} -1 \\ 0 \\ -4 \\ 2 \end{pmatrix}$. Cuantas rectas hay con estas condiciones?

52. Encuentre una ecuación de una recta que contenga el punto P del item anterior y sea paralela al vector $\mathbf{v} = \begin{pmatrix} -1 \\ 0 \\ -4 \\ 2 \end{pmatrix}$. Cuantas rectas hay con estas condiciones?

53. Encuentre una ecuación de una recta que contenga el punto P del Ejercicio 50 y sea ortogonal a la recta $\mathcal{L}$ del Ejercicio 49. Cuántas rectas hay con estas condiciones?

54. Encuentre dos puntos del plano $\mathcal{P}$ con ecuaciones paramétricas: $\begin{cases} x_1 = -3 - 2t - s \\ x_2 = s \\ x_3 = -t - 1 \\ x_4 = 0 \\ x_5 = t + s \end{cases}$

55. Demuestre el Teorema 14.

56. Encuentre una ecuación de un plano que contenga los puntos

$$P = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 5 \\ -3 \end{pmatrix}, \quad Q = \begin{pmatrix} 2 \\ 2 \\ -3 \\ 5 \\ 0 \end{pmatrix}, \quad R = \begin{pmatrix} -1 \\ 2 \\ 0 \\ -3 \\ 5 \end{pmatrix}.$$

Cuantos planos hay con estas condiciones?

57. Encuentre una ecuación de un plano que contenga el punto P del Ejercicio 55 y sea paralelo al plano $\mathcal{P}$ del Ejercicio 53. Cuántos planos hay con estas condiciones?

58. Encuentre una ecuación de un plano que contenga el punto P del Ejercicio 50 y sea ortogonal a la recta $\mathcal{L}$ del Ejercicio 49. Cuántos planos hay con estas condiciones?

59. Encuentre una ecuación de un plano que contenga el punto P del Ejercicio 50 y contenga la recta $\mathcal{L}$ del Ejercicio 49. Cuántos hay con estas condiciones?

60. Encuentre una ecuación de un hiperplano que contenga los puntos $P = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 5 \end{pmatrix}$, $Q = \begin{pmatrix} 2 \\ 2 \\ -3 \\ 5 \end{pmatrix}$ y $R = \begin{pmatrix} -1 \\ 2 \\ 0 \\ -3 \end{pmatrix}$(AYUDA: llame $\mathbf{n} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ al vector normal y recuerde que $\mathbf{n}$ debe ser ortogonal a $\overline{PQ}$ y $\overline{PR}$). Cuántos hiperplanos hay con estas condiciones?

61. Encuentre una ecuación de un hiperplano que contenga, además de los puntos P, Q y R del ejercicio anterior, el punto $S = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$(AYUDA: llame $\mathbf{n} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ al vector normal y recuerde que $\mathbf{n}$ debe ser ortogonal a $\overline{PQ}$, $\overline{PR}$ y $\overline{PS}$). Cuántos hiperplanos hay con estas condiciones?

62. Encuentre una ecuación de un hiperplano que contenga el punto P del Ejercicio 50 y sea paralelo al hiperplano $\mathcal{H}:\ 3x_1 - 2x_3 - 5 = -x_4$. Cuántos hiperplanos hay con estas condiciones?

63. Encuentre una ecuación de un hiperplano que contenga el punto P del Ejercicio 50 y sea ortogonal al hiperplano $\mathcal{H}:\ 3x_1 - 2x_3 - 5 = -x_4$. Cuántos hiperplanos hay con estas condiciones?

64. Determine si la recta $\mathcal{L}:\ \dfrac{x+5}{-3} = 1 + z = \dfrac{w}{2},\ \ y = 0$ intercepta a cada uno de los siguientes planos

a) $\mathcal{P}: \begin{cases} x_1 = 1 - 2t \\ x_2 = 2t - 4s \\ x_3 = -3t + 7s \\ x_4 = 4t - 6s \end{cases}$ b) $\mathcal{P}: \begin{cases} x_1 = -t - s \\ x_2 = 5t - 6s \\ x_3 = -4t + 4s \\ x_4 = -4 \end{cases}$

En caso afirmativo, encuentre la intersección.

65. Demuestre la segunda parte del Teorema 19.

Para los Ejercicios 65 a 68, sean $\mathcal{L}_1,\ \mathcal{L}_2,\ \mathcal{L}_3$ y $\mathcal{P}_1,\ \mathcal{P}_2,\ \mathcal{P}_3$ los siguientes rectas y planos de R^3 y P un punto de R^3.

$\mathcal{L}_1: \begin{cases} x = -t \\ y = 1 + 2t \\ z = -2 + 3t \end{cases}$ $\mathcal{P}_1: \begin{cases} x_1 = -3 - 2t \\ x_2 = 5t - 6s \\ x_3 = 1 - 4t + 4s \end{cases}$ $P = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$

$\mathcal{L}_2:\ \dfrac{x+5}{6} = \dfrac{z-1}{2}\quad y = -1$ $\mathcal{P}_2:\ x - 4y + 3z - 7 = 0$

$\mathcal{L}_3: \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} + t \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$ $\mathcal{P}_3: \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} + t \begin{pmatrix} 3 \\ -3 \\ -5 \end{pmatrix} + s \begin{pmatrix} 6 \\ 0 \\ -2 \end{pmatrix}$

66. Encuentre

a) un punto de cada recta y cada plano.

b) un vector director de cada recta.

c) los otros dos tipos de ecuaciones de las rectas.

d) dos vectores directores de cada plano.

e) un vector normal de cada plano.

f) los otros dos tipos de ecuaciones de los planos.

67. Determine

a) cuales rectas son ortogonales y cuales son paralelas.

b) cuales planos son ortogonales y cuales son paralelos.

c) cual de las rectas es ortogonal al plano $\mathcal{P}_1$.

d) cual de las rectas es paralela al plano $\mathcal{P}_2$.

e) cual de las rectas corta al plano $\mathcal{P}_3$.

f) cual de las rectas está contenida en el plano $\mathcal{P}_1$.

68. Encuentre

a) una recta paralela a la recta $\mathcal{L}_1$ que pase por el origen. Existe otra?

b) una recta ortogonal a la recta $\mathcal{L}_2$ que corte a la recta $\mathcal{L}_3$. Existe otra?

c) un plano que contenga la recta $\mathcal{L}_1$. Existe otro?

d) un plano paralelo a la recta $\mathcal{L}_2$ que pase por el origen. Existe otro?

e) un plano ortogonal a la recta $\mathcal{L}_3$ que contenga a una de las otras dos rectas. Existe otro?

f) un plano paralelo al plano $\mathcal{P}_1$ que pase por P. Existe otro?

g) un plano ortogonal al plano $\mathcal{P}_2$ que contenga a la recta $\mathcal{L}_1$. Existe otro?

69. Calcule

a) la distancia del punto P a la recta $\mathcal{L}_1$ (AYUDA: Calcule la norma de la componente vectorial del vector $\overline{PQ}$ ortogonal a $\mathbf{v}$, siendo Q un punto de la recta y $\mathbf{v}$ un vector director de la recta).

b) la distancia de un punto P al plano $\mathcal{P}_1$ (AYUDA: Calcule la norma de la proyección ortogonal del vector $\overline{PQ}$ sobre $\mathbf{n}$, siendo Q un punto del plano y $\mathbf{n}$ un vector normal del plano).

c) la distancia entre las recta $\mathcal{L}_1$ y $\mathcal{L}_2$ (AYUDA: Si las rectas son paralelas, calcule la distancia de un punto cualquiera de una de las rectas a la otra [ver a)]. En caso contrario, calcule la norma de la proyección ortogonal del vector $\overline{PQ}$ sobre $\mathbf{n}$, siendo P un punto de la recta $\mathcal{L}_1$, Q un punto de la recta $\mathcal{L}_2$ y $\mathbf{n}$ un vector ortogonal a las dos rectas).

d) la distancia de la recta $\mathcal{L}_1$ al plano $\mathcal{P}_1$ (AYUDA: Si la recta es paralela al plano, calcule la distancia de un punto cualquiera de la recta $\mathcal{L}_1$ al plano [ver b)]. En caso contrario, la distancia es cero).

e) la distancia entre los planos $\mathcal{P}_1$ y $\mathcal{P}_2$ (AYUDA: Si los planos son paralelos, calcule la distancia de un punto cualquiera de uno de los planos al otro plano [ver b)]. En caso contrario, la distancia es cero).

70. Hallar una fórmula general para calcular la distancia

a) entre un punto P a una recta $\mathcal{L}$

b) entre un punto P a un plano $\mathcal{P}$.

c) entre dos rectas $\mathcal{L}_1$ y $\mathcal{L}_2$.

d) entre una recta $\mathcal{L}$ y un plano $\mathcal{P}$.

e) entre dos planos $\mathcal{P}_1$ y $\mathcal{P}_2$.

71. Complete la demostración del Teorema 16.

Para los ejercicios 71 y 72, sean

$$O = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \; P = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}, \; Q = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \; R = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix},$$

puntos de R^3 y $\mathbf{u} = \overline{OP}$, $\mathbf{v} = \overline{OQ}$ y $\mathbf{w} = \overline{OR}$.

72. Determine cuales de las siguientes expresiones están bien definidas y, en caso positivo, haga los cálculos indicados.

a) $(2\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$.
b) $(\mathbf{u} \times 2\mathbf{v}) \cdot (-3\mathbf{w})$.
c) $(\mathbf{u} \cdot 5\mathbf{v}) \times \mathbf{w}$.
d) $(\mathbf{u} \times 2\mathbf{u}) \cdot (-3\mathbf{w})$.
e) $-3(\mathbf{v} \times 2\mathbf{u}) \cdot \mathbf{w}$.
f) $(\mathbf{u} \times 2\mathbf{v}) \cdot (-3\mathbf{u})$.

73. Calcule

a) el área de un paralelogramo en que P, Q y R son tres de sus vértices. Cuantos paralelogramos con estas condiciones existen?, Cuales son sus áreas?

b) el volumen del paralelepípedo cuyos lados son los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$. Son coplanares estos vectores?

Capítulo 3

MATRICES

3.1. Introducción

En los capítulos anteriores, utilizando la noción de matriz, simplificamos la representación de problemas como los sistema de ecuaciones lineales y de conceptos como el de combinación lineal. De paso, esta representación también nos simplificó el análisis y la solución de problemas que involucran un sistema de ecuaciones lineales como la determinación de independencia lineal de un conjunto de vectores y la decisión de si un vector es o no combinación lineal de un conjunto de vectores dado, tomando como herramienta de cálculo una versión matricial del algoritmo de Eliminación de Gauss (el cual calcula la forma escalonada equivalente de una matriz) presentado en el primer capítulo. Sin restarle importancia a todo lo anterior, debemos decir que las matrices son importantes en si mismas y el estudio detallado de sus propiedades algebraicas enriquece su aplicación tanto en la matemática misma como en las demás ciencias y profesiones del mundo actual.

En este capítulo, retomaremos la definición dada de matriz para estudiar las operaciones algebraicas entre ellas (suma, multiplicación por escalar, producto, transpuesta y determinantes), sus propiedades básicas y los principales tipos de matrices que brindan las aplicaciones más comunes de este objeto matemático. Como una aplicación inmediata de la teoría de matrices, presentaremos otra formulación matricial del algoritmo de eliminación de Gauss y su estrecha relación con la factorización LU de una matriz, la cual es la base del desarrollo de los algoritmos computacionales modernos para resolver eficientemente un sistema de ecuaciones lineales y que son objeto de estudio en cursos como Análisis Numérico o Álgebra lineal Numérica.

3.2. Definición y Tipo de Matrices

Definición 1 [*Matriz*]. Una *matriz* es un arreglo rectangular de números reales[1], llamados *componentes* o *elementos* de la matriz, de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

La *i-ésima fila* y la *j-ésima columna* de la matriz A son

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

[1]Aunque los números podrían ser complejos (matrices complejas), en este texto nos limitamos a las matrices de componentes reales (matrices reales)

respectivamente. El *orden o tamaño de la matriz* lo determinamos por el número de filas seguido del número de columnas. Así, decimos que el tamaño de una matriz es $m \times n$, si tiene m filas y n columnas; por simplicidad, cuando $m = n$, decimos que la matriz es de orden n . La componente a_{ij} ((i,j)-ésima componente) se encuentra en la fila i y columna j de la matriz A. También podemos denotar la matriz A como $\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$, donde cada $\mathbf{a}_i$ (columna i de la matriz A) es un vector de R^m ó como (a_{ij}), donde a_{ij} (componente (i,j) de la matriz A) es un número real.

Ejemplo 1. Sean

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 5 & 2 & -3 \\ -2 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1/3 & -1 & 4 & 1 \\ 5 & \sqrt{2} & -3 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 4 \\ 5 \\ -2 \\ 0 \\ 1 \end{pmatrix},$$

$$D = \begin{pmatrix} 1 & -1 & 1 \end{pmatrix} \quad \text{y} \quad E = (e_{ij}),\ e_{ij} = (-1)^{i+j},\ i,j = 1,2,\ldots n.$$

La matriz A es de orden 3×3, B es una matriz 2×4, C es una matriz 5×1 (en general, un vector de R^m es una matriz $m \times 1$), D es una matriz 1×3 (algunos autores llaman *vector fila* de R^n a las matrices $1 \times n$) y E es una matriz $n \times n$ o simplemente de orden n. □

La *diagonal de una matriz* está formada por las componentes a_{ij} , con $i = j$; es decir, por las componentes a_{ii}. Así, podemos ver que en el Ejemplo 1, las diagonales de las matrices A, B, C y D son $\begin{pmatrix} 1 & 2 & 1 \end{pmatrix}$, $\begin{pmatrix} 1/3 & \sqrt{2} \end{pmatrix}$, $\begin{pmatrix} 4 \end{pmatrix}$ y $\begin{pmatrix} 1 \end{pmatrix}$, respectivamente.

Cuando las componentes arriba de la diagonal son ceros ($a_{ij} = 0$ para $i < j$), tenemos una *matriz triangular inferior*; si las componentes abajo de la diagonal son ceros ($a_{ij} = 0$ para $i > j$), tenemos una *matriz triangular superior*. Ahora, si el número de filas de una matriz coincide con el número de columnas ($m = n$), tenemos una *matriz cuadrada*. De las matrices cuadradas, si tanto los elementos arriba de la diagonal, como los elementos abajo de la diagonal son ceros ($a_{ij} = 0$ para $i \neq j$), tenemos una *matriz diagonal* y si la matriz diagonal tiene todos los elementos de la diagonal iguales ($a_{ij} = 0$ para $i \neq j$ y $a_{ii} = \alpha$), tenemos una *matriz escalar*. La matriz escalar más importante es la que tiene todos los elementos de la diagonal iguales a 1 ($a_{ij} = 0$ para $i \neq j$ y $a_{ii} = 1$). Por las propiedades que demostraremos más adelante en el *producto de matrices*, a esta matriz la llamamos *matriz identidad o idéntica* y la denotamos como I ó I_n cuando queramos explicitar que su orden es n. Observemos que

$$I_n = [\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots,\ \mathbf{e}_n],$$

donde los $\mathbf{e}_i$ son los vectores canónicos de R^n definidos en el Ejemplo 2 del Capítulo 2.

Ejemplo 2. Sean

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ -2 & 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1/3 & -1 & 4 & 1 \\ 0 & -1 & 4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix} \quad e \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Así, A, C, D e I son matrices triangulares inferiores; B, C, D e I son matrices triangulares superiores; C, D e I son matrices diagonales; D e I son matrices escalares; e I es la matriz idéntica de orden 3 o I_3. □

Otra matriz importante (por sus propiedades respecto de las operaciones que definiremos más adelante) es la *matriz cero o matriz nula*, la cual tiene todas sus componentes iguales a cero y la denotaremos como O ó $O_{m\times n}$ cuando queramos explicitar que su orden es $m \times n$.

De otro lado, al igual que para vectores, diremos que dos matrices son *iguales*, cuando todas sus componentes respectivas son iguales, para lo cual se requiere que sus tamaños también sean iguales.

Ejemplo 3. Consideremos las matrices

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 2 & -1 & 3 \\ 0 & -2 & 5 \end{pmatrix}.$$

Para que las matrices A y B sean iguales, se debe tener $a = 2$, $b = -1$, $c = 3$, $d = 0$, $e = -2$ y $f = 5$. □

Ejemplo 4. Consideremos las matrices

$$C = \begin{pmatrix} -1 & 3 & 5 \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix}.$$

Estas matrices son diferentes, ya que sus tamaños son diferentes. □

3.3. Suma y Producto por escalar de Matrices

Al igual que con los escalares y los vectores, dadas unas matrices, es deseable obtener otras a partir de ellas y para esto, podemos generalizar las definiciones de la suma de vectores a la suma de matrices y del producto por escalar de vectores al producto por escalar de matrices.

Definición 2 [*Suma de matrices*]. Definimos la *suma entre dos matrices* de igual tamaño A y B, como la matriz $A + B$, cuyas componentes son la suma de las componentes respectivas de las matrices A y B. En otras palabras, dadas las matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix},$$

se define la suma de A y B por

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Definición 3 [*Producto por escalar de matrices*]. Definimos el producto de una matriz A por un número real λ (escalar), como la matriz λA, cuyas componentes son el producto de λ por las componentes respectivas de A. En otras palabras, dado el escalar $\lambda \in R$ y la matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

definimos el producto de A por el escalar λ por

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

Ejemplo 5. Dadas las matrices

$$A = \begin{pmatrix} 1 & 2 & -5 \\ -1 & -3 & 0 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} -3 & 2 & 1 \\ 0 & 1 & -2 \end{pmatrix},$$

tenemos que

$$A + B = \begin{pmatrix} -2 & 4 & -4 \\ -1 & -2 & -2 \end{pmatrix} \quad \text{y} \quad -2A = \begin{pmatrix} -2 & -4 & 10 \\ 2 & 6 & 0 \end{pmatrix}.$$

□

El producto por escalar $(-1)A$, lo denotamos simplemente como $-A$ y así, la resta $A - B$, la obtenemos de $A + (-1)B$.

Al igual que para vectores, la suma y el producto por escalar de matrices tienen propiedades muy similares a las que poseen los números reales. Veamos cuales son las propiedades que se satisfacen para estas dos operaciones entre matrices.

Teorema 1 [*Propiedades algebraicas de la suma y la multiplicación por escalar de matrices*].
Sean A, B y C matrices de tamaño $m \times n$ y sean α y β dos números reales. Entonces se cumplen las siguientes proposiciones:

1. $A + B$ es también una matriz $m \times n$. *Ley clausurativa para la suma*
2. $A + B = B + A$. *Ley conmutativa para la suma*
3. $(A + B) + C = A + (B + C)$. *Ley asociativa para la suma*
4. Existe una única matriz Z de tamaño $m \times n$ tal que $A + Z = Z + A = A$ $(Z = O)$. *Ley modulativa para la suma*
5. Para cada matriz A, existe una única matriz P de tamaño $m \times n$ tal que $A + P = P + A = O$ $(P = -A)$. *Existencia del opuesto para la suma*
6. λA es también una matriz $m \times n$. *Ley clausurativa para el producto por escalar*
7. $\alpha(A + B) = \alpha A + \alpha B$. *Ley distributiva del producto por escalar respecto a la suma de matrices*
8. $(\alpha + \beta)A = \alpha A + \beta A$. *Ley distributiva del producto por escalar respecto a la suma de escalares*
9. $(\alpha\beta)A = \alpha(\beta A) = \beta(\alpha A)$. *Ley asociativa respecto al producto por escalar*
10. $1A = A$. *Ley modulativa para el producto por escalar*
11. $0A = O$.
12. $\alpha O = O$.
13. $\alpha A = O$, si y solo si, $\alpha = 0$ ó $A = O$.

La demostración de estas propiedades es muy similar a la de las propiedades de estas operaciones entre vectores (Teorema 1, Capítulo 2) y la dejamos como ejercicios para el lector. □

Estas propiedades nos permiten resolver ecuaciones sencillas entre matrices, simplificar expresiones matriciales y proporcionan diferentes formas para hacer un mismo cálculo lo cual permite elegir la forma más eficiente de realizarlo.

Ejemplo 6. Dadas las matrices $A = \begin{pmatrix} 0 & -6 \\ 3 & 0 \\ -1 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} -2 & 3 \\ 4 & 1 \\ 0 & -1 \end{pmatrix}$, determine la matriz X tal que

$$3X - 2A + B = 4B.$$

Sumando el opuesto de las matrices $-2A$ y B en ambos lados de la ecuación, obtenemos

$$\begin{aligned}
3X - 2A + B + 2A - B &= 4B + 2A - B & \\
3X + 2A - 2A + B - B &= 4B - B + 2A & \text{Propiedad 2, Teorema 1.}\\
3X + (2A - 2A) + (B - B) &= (4B - B) + 2A & \text{Propiedad 3, Teorema 1.}\\
3X + O + O &= 3B + 2A & \text{Propiedad 5, Teorema 1.}\\
3X &= 3B + 2A & \text{Propiedad 4, Teorema 1.}
\end{aligned}$$

Multiplicando por el escalar $\frac{1}{3}$ ambos lados de la última ecuación, tenemos

$$\begin{aligned}
\frac{1}{3}(3X) &= \frac{1}{3}(3B + 2A)\\
\left(\frac{1}{3}3\right)X &= \frac{1}{3}(3B) + \frac{1}{3}(2A)\\
X &= B + \frac{2}{3}A.
\end{aligned}$$

Así que

$$\begin{aligned}
X &= \begin{pmatrix} -2 & 3 \\ 4 & 1 \\ 0 & -1 \end{pmatrix} + \frac{2}{3}\begin{pmatrix} 0 & -6 \\ 3 & 0 \\ -1 & 3 \end{pmatrix}\\
&= \begin{pmatrix} -2 & -1 \\ 6 & 1 \\ -2/3 & 1 \end{pmatrix}.
\end{aligned}$$

□

Observemos que $cA + cB = c(A + B)$, pero la segunda expresión $(c(A + B))$ requiere menos multiplicaciones que la primera, por lo tanto, es más eficiente para realizar este cálculo.

3.4. Producto de Matrices

Otra operación que uno espera que se pueda definir entre matrices es la multiplicación o producto. Sin embargo, la definición no es la que, tal vez, uno se imagina. La definición que daremos tiene significado en la aplicación de la teoría de matrices a problemas reales y está relacionada con el producto $A\mathbf{x}$ definido en el capítulo anterior.

Definición 4 [*Producto de matrices*]. Dadas las matrices A, de tamaño $m \times n$, y B, de tamaño $n \times k$, se define el producto de A por B, como la matriz AB, cuyas columnas son $A\mathbf{b}_1\ A\mathbf{b}_2\ \cdots\ A\mathbf{b}_k$, donde $\mathbf{b}_1\ \mathbf{b}_2\ \cdots \mathbf{b}_k$ son las columnas de B; en otras palabras, se define AB, el producto de A por $B = [\mathbf{b}_1\ \mathbf{b}_2\ \cdots\ \mathbf{b}_k]$, como la matriz de orden $m \times k$ dada por

$$AB = [A\mathbf{b}_1\ A\mathbf{b}_2\ \cdots\ A\mathbf{b}_k].$$

Para enfatizar el orden de los factores en el producto AB, decimos que A *pre-multiplica* a B o que B *post-multiplica* a A. Al producto AA, lo denotaremos A^2 y, en general, a $\underbrace{AA\ldots A}_{n \text{ veces}}$ por A^n, para $n \in N$.

Ejemplo 7. Dadas las matrices $A = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ -1 & 5 \end{pmatrix}$ y $B = \begin{pmatrix} -2 & 3 \\ 4 & 1 \end{pmatrix}$, calculemos AB o pre-multiplequemos a B por A.

Usando la Definición 7 del Capítulo 2, tenemos

$$A\mathbf{b}_1 = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} -2 \\ 4 \end{pmatrix} = -2\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} + 4\begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} -10 \\ -6 \\ 22 \end{pmatrix}$$

$$A\mathbf{b}_2 = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} + 1\begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 9 \\ 2 \end{pmatrix}.$$

Así que,

$$AB = \left[\ A\mathbf{b}_1 \quad A\mathbf{b}_2\ \right] = \begin{pmatrix} -10 & 1 \\ -6 & 9 \\ 22 & 2 \end{pmatrix}.$$

□

Observemos que el producto de matrices sólo es posible si el número de columnas de la primera matriz es igual al número de filas de la segunda matriz. De otro lado, no es difícil ver que la $(i, j)-$ésima componente de AB es la i-ésima componente del vector $A\mathbf{b}_j$, la cual es el resultado del producto escalar de la i-ésima fila de A por la j-ésima columna de B. En efecto, volvamos sobre el ejemplo anterior y verifiquemos que la componente $(3, 1)$ del producto AB efectivamente es el producto escalar de la tercera fila de la matriz A y la primera columna de la matriz B.

$$AB = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ \mathbf{-1} & \mathbf{5} \end{pmatrix}\begin{pmatrix} \mathbf{-2} & 3 \\ \mathbf{4} & 1 \end{pmatrix} = \begin{pmatrix} -10 & 1 \\ -6 & 9 \\ \mathbf{22} & 2 \end{pmatrix}.$$

Resumamos la observación anterior en el siguiente teorema. La demostración queda como ejercicio para el lector, no sin antes anotar que otros autores utilizan este resultado como definición del producto de matrices.

Teorema 2 [*Componentes del producto de matrices*].
Dadas las matrices $A = (a_{ij})$, de tamaño $m \times n$, y $B = (b_{jk})$, de tamaño $n \times p$, si $C = (c_{ik}) = AB$, entonces

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \mathbf{a}_i \cdot \mathbf{b}_k\ ,$$

donde $\mathbf{a}_i$ y $\mathbf{b}_k$ son los vectores de R^n formados por los elementos de la fila i de A y por lo elementos de la columna j de B, respectivamente, para $i = 1,\ 2,\ \ldots\ m$ y $k = 1,\ 2,\ \ldots\ p$. □

Antes de presentar las propiedades algebraicas del producto de matrices, veamos el significado de las operaciones entre matrices presentadas hasta ahora con una situación real en miniatura.

Ejemplo 8. Sean J y N las matrices que resumen la información de las ventas realizadas por Juliana y Nathalie, respectivamente, durante 1 mes (4 semanas).

	Juliana (J)				**Nathalie (N)**			
	Semana				**Semana**			
Producto	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
a	17	3	10	0	40	13	17	23
b	4	5	8	2	8	2	5	0
c	25	15	30	10	3	5	0	6

Sea C la matriz que resume las comisiones (en efectivo y en especie) por artículo vendido durante cada una

de las 4 semanas del mes.

	Comisión por artículo vendido	
Semana	**En efectivo**	**En especie**
	(Miles de $)	(No. de Unidades)
1	12,50	2
2	20,00	3
3	27,50	1
4	9,50	5.

1. El total de las ventas realizadas por las dos vendedoras estará dado por

$$J+N=\begin{pmatrix} 17 & 3 & 10 & 0 \\ 4 & 5 & \mathbf{8} & 2 \\ 25 & 15 & 30 & 10 \end{pmatrix}+\begin{pmatrix} 40 & 13 & 17 & 23 \\ 8 & 2 & \mathbf{5} & 0 \\ 3 & 5 & 0 & 6 \end{pmatrix}=\begin{pmatrix} 57 & 16 & 27 & 23 \\ 12 & 7 & \mathbf{13} & 2 \\ 28 & 20 & 30 & 16 \end{pmatrix},$$

   por ejemplo, entre las dos, vendieron $13 = 8 + 5$ artículos del Producto **b** durante la Semana 3.

2. Si Juliana repite su esquema de ventas durante 6 meses, el total de ventas de Juliana para este período estará dado por

$$6J=6\begin{pmatrix} 17 & \mathbf{3} & 10 & 0 \\ 4 & 5 & 8 & 2 \\ 25 & 15 & 30 & 10 \end{pmatrix}=\begin{pmatrix} 102 & \mathbf{18} & 60 & 0 \\ 24 & 30 & 48 & 12 \\ 150 & 90 & 180 & 60 \end{pmatrix},$$

   por ejemplo, en las condiciones dadas, Juliana vendió $18 = 6 \times 3$ artículos del Producto **a** durante las segundas semanas de estos 6 meses.

3. La distribución, por tipo de producto, de la comisión devengada por Nathalie durante 1 mes estará dado por

$$NC=\begin{pmatrix} 40 & 13 & 17 & 23 \\ 8 & 2 & 5 & 0 \\ \mathbf{3} & \mathbf{5} & \mathbf{0} & \mathbf{6} \end{pmatrix}\begin{pmatrix} \mathbf{12{,}50} & 2 \\ \mathbf{20{,}00} & 3 \\ \mathbf{27{,}50} & 1 \\ \mathbf{9{,}50} & 5 \end{pmatrix}=\begin{pmatrix} 1.436{,}50 & 251 \\ 277{,}50 & 27 \\ \mathbf{194{,}50} & 51 \end{pmatrix},$$

   por ejemplo, Nathalie obtuvo una comisión, en efectivo, de $\$194{,}500 = 3 \times \$12{,}50 + 5 \times \$20{,}00 + 0 \times \$27{,}50 + F_{\times}\$9{,}50$ por lo vendido del Producto **c** durante el mes en cuestión. □

En el siguiente teorema, se encuentra un resumen de las propiedades algebraicas del producto entre matrices.

Teorema 3 [*Propiedades algebraicas del producto de matrices*].
Sea A una matriz de tamaño $m \times n$, α un número real cualquiera y B y C matrices tales que los productos indicados se pueden efectuar. Entonces se cumplen las siguientes proposiciones:

1. $(AB)C = A(BC)$. *Ley asociativa para el producto*

2. $A(B + C) = AB + AC$. *Ley distributiva del producto de matrices a izquierda respecto de la suma de matrices*

3. $(A + B)C = AC + BC$. *Ley distributiva del producto de matrices a derecha respecto de la suma de matrices*

4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

5. Existe una única matriz P de orden m y una única matriz Q de orden n tales que $PA = AQ = A$. ($P = I_m$ y $Q = I_n$) *Ley modulativa del producto*

6. $O_{km}A = O_{kn}$ y $AO_{nk} = O_{mk}$.

Demostración: Probemos las Propiedades 1 y 4. Dejamos las demás como ejercicio para el lector. Supongamos que

$$C = [\ \mathbf{c}_1 \quad \mathbf{c}_2 \quad \cdots \quad \mathbf{c}_p\ ] \quad , \quad B = [\ \mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_k\ ] \quad \text{y} \quad \mathbf{c}_i = \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ki} \end{pmatrix},$$

1. Observemos que

$$\begin{aligned} (AB)\mathbf{c}_i &= [\ A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_k\ ]\ \mathbf{c}_i \\ &= c_{1i}(A\mathbf{b}_1) + c_{2i}(A\mathbf{b}_2) + \cdots + c_{ki}(A\mathbf{b}_k) \\ &= A(c_{1i}\mathbf{b}_1) + A(c_{2i}\mathbf{b}_2) + \cdots + A(c_{ki}\mathbf{b}_k) \\ &= A(c_{1i}\mathbf{b}_1 + c_{2i}\mathbf{b}_2 + \cdots + c_{ki}\mathbf{b}_k) \\ &= A(B\mathbf{c}_i). \end{aligned}$$

Así que,

$$\begin{aligned} (AB)C &= [\ (AB)\mathbf{c}_1 \quad (AB)\mathbf{c}_2 \quad \cdots \quad (AB)\mathbf{c}_p\ ] \\ &= [\ A(B\mathbf{c}_1) \quad A(B\mathbf{c}_2) \quad \cdots \quad A(B\mathbf{c}_p)\ ] \\ &= A\ [\ B\mathbf{c}_1 \quad B\mathbf{c}_2 \quad \cdots \quad B\mathbf{c}_p\ ] \\ &= A(BC). \end{aligned}$$

4. Ahora,

$$\begin{aligned} \alpha(AB) &= \alpha\ [\ A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_k\ ] \\ &= [\ \alpha(A\mathbf{b}_1) \quad \alpha(A\mathbf{b}_2) \quad \cdots \quad \alpha(A\mathbf{b}_k)\ ] \\ &= [\ (\alpha A)\mathbf{b}_1 \quad (\alpha A)\mathbf{b}_2 \quad \cdots \quad (\alpha A)\mathbf{b}_k\ ] \\ &= (\alpha A)[\ \mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_k\ ] \\ &= (\alpha A)B. \end{aligned}$$

y de manera similar, podemos ver que $\alpha(AB) = A(\alpha B)$ □

Estas propiedades permiten simplificar cálculos. Dos casos fundamentales son los siguientes.

1. ***El cálculo de*** A^n. Por ejemplo, para calcular A^8, tendríamos que efectuar 7 productos de matrices:

$$A^8 = ((((((AA)A)A)A)A)A)A$$

lo que, utilizando la propiedad asociativa de la multiplicación, lo podemos reducir a 3 productos

$$A^2 = AA \quad , \quad A^4 = A^2A^2 \quad \text{y} \quad A^8 = A^4A^4,$$

al asociar el producto de las matrices de la siguiente manera

$$A^8 = [(AA)(AA)][(AA)(AA)].$$

2. ***El cálculo de*** $A^n\mathbf{x}$. Por ejemplo, para calcular

$$A^4\mathbf{x} = (AAAA)\mathbf{x} = [(AA)(AA)]\mathbf{x},$$

es más eficiente calcular

$$A(A(A(A\mathbf{x})));$$

ya que se requieren menos operaciones al multiplicar una matriz por un vector, que al multiplicar dos matrices.

La propiedad asociativa del producto nos permite formular las siguientes propiedades, las cuales son útiles en la simplificación de cálculos, como lo vimos en el primero de los anteriores casos fundamentales.

Teorema 4 [*Propiedades algebraicas de la potenciación de matrices*].
Sean A y B matrices de orden $n \times n$, y r, s números enteros no negativos, entonces

1. $A^r A^s = A^{r+s}$.
2. $(A^r)^s = A^{rs}$.
3. $(AB)^r = A^r B^r$, cuando A y B conmutan ($AB = BA$).

Demostración: La dejamos como ejercicio para el lector. □

Es importante que observemos que la propiedad conmutativa de la multiplicación no siempre se tiene:

1. El producto AB puede estar definido, mientras que BA no. Por ejemplo, si el tamaño de A es $m \times n$ y el de B es $n \times p$ con $m \neq p$.
2. Los dos productos pueden estar bien definidos, pero AB tener un tamaño y BA otro. Por ejemplo, si el tamaño de A es $m \times n$ y el de B es $n \times m$ con $m \neq n$, el tamaño de AB es $m \times m$ y el de la matriz BA es $n \times n$.
3. Si A y B son matrices cuadradas de igual tamaño, tenemos que AB y BA están bien definidas y son de igual tamaño, pero en general $AB \neq BA$.

   Ejemplo 9. Observemos el producto de las siguientes matrices, en diferente orden.

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

   y

$$BA = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}.$$

□

Cuando $AB = BA$, se dice que las *matrices conmutan*.

Igualmente hay otras propiedades del producto que tenemos en los números reales, que entre matrices no tenemos. Esto se debe fundamentalmente a la ausencia de lo que en los números reales sería el inverso multiplicativo. Veamos algunas de ellas:

1. $AB = O$ no implica que A o B sean la matriz O, como vimos en la primera parte del ejemplo anterior.
2. $CA = CB$ (ó $AC = BC$) no implica que $A = B$.

   Ejemplo 10. Observemos que el producto de las siguientes matrices tienen el mismo resultado, y aunque una de ellas es la misma, las otras no son iguales.

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -2 \\ 0 & 2 \end{pmatrix}.$$

□

3. $A^2 = I$ no implica que $A = \pm I$.

   Ejemplo 11. Observemos que el producto de una matriz por si misma da la matriz idéntica; sin embargo, la matriz no es la idéntica ni el opuesto de la idéntica

$$\begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

□

3.5. Matrices invertibles

Volviendo sobre el comentario de la ausencia de inverso multiplicativo, podemos ver que no para todas las matrices A existe otra matriz B tal que $AB = I$.

Ejemplo 12. Veamos que si $A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$, no existe una matriz $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, tal que $AB = I$.

$$\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a-c & b-d \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

□

A las matrices que tienen inverso multiplicativo, les llamamos de manera especial.

Definición 5 [*Matriz invertible*][2]. Se dice que la matriz A de tamaño $n \times n$ es *invertible*, si y solo si, existe una matriz B tal que

$$AB = BA = I.$$

A esta matriz B, la llamamos *inversa* de A.

Surgen, entonces, las siguientes preguntas. Es la matriz inversa única? Cómo determinar que una matriz es invertible? (es necesario encontrar explícitamente una inversa?) Cómo calcular una inversa de una matriz?

Comencemos a contestar estas preguntas, suponiendo que existen dos matrices inversas de A. Sean B y C matrices inversas de A. Así que

$$AB = BA = I \quad \text{y} \quad AC = CA = I$$

De modo que, si multiplicamos ambos lados de la igualdad $AB = I$ por la matriz C y usamos las propiedades del producto de matrices, tenemos

$$\begin{aligned} C(AB) &= CI \\ (CA)B &= C \\ IB &= C \\ B &= C. \end{aligned}$$

De esta forma, hemos demostrado el siguiente teorema.

Teorema 5 [*Unicidad de la inversa*].
Si A es una matriz invertible, su inversa es única. □

Con este resultado, cuando la matriz A es invertible, podemos hablar de ***la*** inversa de A y la denotamos por A^{-1}.

Ejemplo 13. Verifiquemos que $B = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & -1 \\ 3 & 1 & 0 \end{pmatrix}$ es la matriz inversa de $A = \frac{1}{4}\begin{pmatrix} 1 & 0 & 1 \\ -3 & 0 & 1 \\ 2 & -4 & 2 \end{pmatrix}$.

Por multiplicación directa, podemos ver que

$$AB = \frac{1}{4}\begin{pmatrix} 1 & 0 & 1 \\ -3 & 0 & 1 \\ 2 & -4 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & -1 \\ 3 & 1 & 0 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

De la misma forma, podemos verificar que $BA = I_3$, por lo tanto, B es la matriz inversa de A; es decir,

$$B = A^{-1}.$$

□

[2]Otro nombre que comunmente reciben las matrices invertibles es el de *matrices no singulares*.

Veamos ahora, cómo determinar si una matriz es invertible. Sabemos que si una matriz A, de tamaño $n \times n$, es invertible, su inversa, A^{-1}, también tiene tamaño $n \times n$.

Asi que debemos verificar si existe una matriz $B = [\ \mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n\ ]$, tal que

$$AB = [\ A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_n\ ] = I = [\ \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n\ ].$$

En otras palabras, tenemos que determinar si los sistemas

$$A\mathbf{b}_1 = \mathbf{e}_1,\ A\mathbf{b}_2 = \mathbf{e}_2,\ \cdots,\ A\mathbf{b}_n = \mathbf{e}_n$$

tienen solución. Para esto, podemos escalonar las matrices

$$[A : \mathbf{e}_1]\ ,\ [A : \mathbf{e}_2]\ , \cdots,\ [A : \mathbf{e}_n]\,,$$

y determinar la consistencia de los sistemas de ecuaciones lineales asociados, lo que se garantiza si la forma escalonada de A tiene n pivotes. Observemos que, para determinar si A es una matriz invertible, no es necesario resolver los anteriores sistemas de ecuaciones lineales. De ser necesario calcular A^{-1}, podríamos resolver simultáneamente los anteriores sistemas mediante el Algoritmo *Eliminación de Gauss + Sustitución hacia atrás*[3] aplicado a la matriz aumentada conjunta

$$[A : \mathbf{e}_1\ \mathbf{e}_2,\ \cdots,\ \mathbf{e}_n] = [A : I],$$

como vimos al final del Capítulo 1.

Ejemplo 14. Determinemos si $A = \begin{pmatrix} 0{,}25 & -0{,}25 & 0 \\ 0{,}50 & 0 & -0{,}25 \\ 0{,}75 & 0{,}25 & 0 \end{pmatrix}$ es invertible.

Aplicando sucesivamente las operaciones elementales $F_2 - 2F_1 \to F_2$, $F_3 - 3F_1 \to F_3$ y $F_3 - 2F_2 \to F_3$ a la matriz A (Eliminación de Gauss), obtenemos la matriz escalonada equivalente

$$\begin{pmatrix} 0{,}25 & -0{,}25 & 0 \\ 0 & 0{,}50 & -0{,}25 \\ 0 & 0 & 0{,}50 \end{pmatrix}.$$

Como cada fila tiene pivote, los sistemas $A\mathbf{b}_1 = \mathbf{e}_1$, $A\mathbf{b}_2 = \mathbf{e}_2$ y $A\mathbf{b}_3 = \mathbf{e}_3$ tienen solución; es decir, la matriz inversa de A existe, por lo tanto A es una matriz invertible. $\square$

Ejemplo 15. Encontremos la inversa de la matriz A del ejemplo anterior.

Como observamos antes, para calcular A^{-1}, debemos resolver simultáneamente los sistemas de ecuaciones lineales

$$A\mathbf{b}_1 = \mathbf{e}_1, \quad A\mathbf{b}_2 = \mathbf{e}_2, \quad \text{y} \quad A\mathbf{b}_3 = \mathbf{e}_3,$$

donde $\mathbf{e}_1, \mathbf{e}_2$ y $\mathbf{e}_3$ son los vectores canónicos de R^3. Para ello, apliquemos el Algoritmo *Eliminación de Gauss + Sustitución hacia atrás*, separando sus dos etapas. Primero, aplicando sucesivamente las mismas operaciones elementales que en el ejemplo anterior a la matriz aumentada conjunta $[A : I_3]$ (Eliminación de Gauss), obtenemos

$$\left(\begin{array}{ccc|ccc} 0{,}25 & -0{,}25 & 0 & 1 & 0 & 0 \\ 0 & 0{,}50 & -0{,}25 & -2 & 1 & 0 \\ 0 & 0 & 0{,}50 & 1 & -2 & 1 \end{array}\right).$$

Segundo, apliquemos sucesivamente las operaciones elementales $2F_3 \to F_3$, $F_2 + 0{,}25F_3 \to F_2$, $2F_2 \to F_2$, $F_1 + 0{,}25F_2 \to F_1$ y $4F_1 \to F_1$ a la anterior matriz (Sustitución hacia atrás), para obtener

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -3 & 0 & 1 \\ 0 & 0 & 1 & 2 & -4 & 2 \end{array}\right).$$

[3]El Algoritmo de Gauss-Jordan que, en programación no paralela, es menos eficiente, también podría usarse para resolver simultáneamente estos sistemas.

Como vimos al final del Capítulo 1, cada una de las tres columnas de la derecha en la matriz anterior es la solución de cada uno de los tres sistemas de ecuaciones lineales planteados, respectivamente; por lo tanto,

$$A^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ -3 & 0 & 1 \\ 2 & -4 & 2 \end{pmatrix}.$$

□

Nuevamente, notemos que si se quiere calcular la inversa de una matriz A, generalmente, es necesario resolver los sistemas $[A : \mathbf{e}_1]$, $[A : \mathbf{e}_2]$, $\cdots$, $[A : \mathbf{e}_n]$; y que para determinar si una matriz es invertible, no es necesario calcular su inversa, puesto que es suficiente con saber que su inversa existe, para lo cual sólo se requiere escalonar la matriz y verificar que la matriz escalonada equivalente tiene n pivotes.

Por suerte, lo que generalmente necesitamos es saber si una matriz es invertible y no saber cuál es su inversa. Incluso, si necesitamos calcular el vector $A^{-1}\mathbf{b}$ para una matriz A y un vector $\mathbf{b}$ dados, es más fácil resolver el sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$, que calcular la matriz inversa de A y luego multiplicar el resultado por el vector b.

Veamos algunas de las propiedades de las matrices invertibles y sus inversas, las cuales son muy importantes tanto en el desarrollo teórico como en la simplificación de expresiones matriciales para realizar cálculos.

Teorema 6 [*Propiedades algebraicas de la inversa de una matriz*].
Sean A y B matrices invertibles de tamaño $n \times n$, λ un escalar diferente de 0 y m un número natural, entonces

1. A^{-1} también es invertible y $(A^{-1})^{-1} = A$.
2. λA también es invertible y $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$.
3. AB también es invertible y $(AB)^{-1} = B^{-1}A^{-1}$(notemos el orden en el producto de las inversas).
4. A^m también es invertible y $(A^m)^{-1} = (A^{-1})^m$.

Demostración:
En la demostración de cada uno de los puntos de este teorema, debemos mostrar que existe una matriz que satisface las dos igualdades de la Definición 5. Para cada caso, demostramos una de ellas y dejamos la otra igualdad como ejercicio para el lector.

1. Para demostrar que A^{-1} es invertible, es suficiente con mostrar que existe una matriz C tal que $A^{-1}C = I$. Pero es claro que si tomamos $C = A$, tenemos $A^{-1}A = I$. Así que A^{-1} es invertible y su inversa (que es única) es A. En otras palabras, $(A^{-1})^{-1} = A$.
2. Veamos que si $C = \frac{1}{\lambda} A^{-1}$, entonces, por el Teorema 3,
$$C(\lambda A) = \left(\frac{1}{\lambda} A^{-1}\right)(\lambda A) = \left(\frac{1}{\lambda}\lambda\right)(A^{-1}A) = I.$$
Así que λA es invertible y $(\lambda A)^{-1} = C = \frac{1}{\lambda} A^{-1}$.
3. De manera similar, vemos que
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$
Así que AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.

4. Igualmente, como A y A^{-1} conmutan, por el Resultado 3 del Teorema 4,

$$A^m(A^{-1})^m = (AA^{-1})^m = I^m = I$$

Asi que A^m también es invertible y $(A^m)^{-1} = (A^{-1})^m$. □

Ahora estamos en capacidad de demostrar uno de los resultados más importantes del Álgebra Lineal relacionado con la caracterización de las matrices invertibles en término de la solución de sistemas de ecuaciones lineales y la independencia lineal de vectores.

Teorema 7 [*Equivalencia de la invertibilidad*].
Sea A una matriz de tamaño $n \times n$. Las siguientes proposiciones son equivalentes.

a) La matriz A es invertible.

b) La única solución de $A\mathbf{x} = \mathbf{b}$ es $\mathbf{x} = A^{-1}\mathbf{b}$.

c) La única solución de $A\mathbf{x} = \mathbf{0}$ es $\mathbf{x} = \mathbf{0}$.

d) Las columnas de la matriz A son linealmente independientes (*l.i.*).

e) Toda matriz escalonada equivalente a A tiene n pivotes.

Demostración:
Para demostrar estas equivalencias, demostremos las siguientes implicaciones

$$a) \Longrightarrow b) \Longrightarrow c) \Longrightarrow d) \Longrightarrow e) \Longrightarrow a)$$

[$a) \Longrightarrow b)$] Supongamos que A es una matriz invertible. Veamos primero que $\mathbf{x} = A^{-1}\mathbf{b}$ es solución del sistema $A\mathbf{x} = \mathbf{b}$ y luego, que esta solución es la única. Para esto, verifiquemos que $\mathbf{x} = A^{-1}\mathbf{b}$ satisface la ecuación $A\mathbf{x} = \mathbf{b}$.

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

Ahora, supongamos que $\mathbf{y}$ es otra solución de $A\mathbf{x} = \mathbf{b}$. Así, $A\mathbf{y} = \mathbf{b}$ y al multiplicar ambos lados de esta igualdad por A^{-1}, tenemos

$$\begin{aligned} A^{-1}(A\mathbf{y}) &= A^{-1}\mathbf{b} \\ (A^{-1}A)\mathbf{y} &= A^{-1}\mathbf{b} \\ \mathbf{y} &= \mathbf{x}. \end{aligned}$$

[$b) \Longrightarrow c)$] Teniendo en cuenta que si $\mathbf{h}$ es solución del sistema homogéneo $A\mathbf{x} = \mathbf{0}$ y $\mathbf{x} = A^{-1}\mathbf{b}$ es solución del sistema $A\mathbf{x} = \mathbf{b}$, por el Corolario 5.1 del Capítulo 2, $\mathbf{x} + \mathbf{h}$ es también solución de $A\mathbf{x} = \mathbf{b}$, como la solución es única, entonces $\mathbf{x} = \mathbf{x} + \mathbf{h}$; por lo tanto, $\mathbf{h} = \mathbf{0}$; es decir, el vector $\mathbf{0}$ es la única solución de $A\mathbf{x} = \mathbf{0}$.

[$c) \Longrightarrow d)$] Si $A\mathbf{x} = \mathbf{0}$ tiene solución única, por el Teorema 6 del Capítulo 2, las columnas de la matriz A son *l.i.*

[$d) \Longrightarrow e)$] Por el Teorema 6 del Capítulo 2, si las columnas de A son *l.i.*, toda forma escalonada equivalente tiene n pivotes.

[$e) \Longrightarrow a)$] Si una forma escalonada equivalente a A tiene n pivotes, todas sus filas tienen pivotes; por lo tanto, el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución para cualquier b; en particular, los sistemas $A\mathbf{x} = \mathbf{e}_i$, para $i = 1, 2, \ldots, n$ tienen solución. Como, por definición, la solución de $A\mathbf{x} = \mathbf{e}_i$ es la columna i de la matriz inversa de A, concluimos que la inversa de A existe, por tanto A es invertible. □

Aunque suene repetitivo, es importante observar que la propiedad *b*) del Teorema 7 no se utiliza para calcular la solución de un sistema de ecuaciones lineales, pues para calcular la matriz inversa habría que resolver los

n sistemas de ecuaciones de la forma $A\mathbf{x} = \mathbf{e}_i$. Como dijimos antes, esta propiedad se debe utilizar en el sentido contrario, si se necesita calcular $A^{-1}\mathbf{b}$, se resuelve el sistema $A\mathbf{x} = \mathbf{b}$, evitándose el cálculo de A^{-1}.

Observemos que el Resultado 3 del Teorema 6 nos da una condición suficiente para que el producto de matrices sea invertible. Veamos que la equivalencia entre las proposiciones a) y c) del Teorema 7 nos permite demostrar que la invertibilidad de los factores también es una condición necesaria para la invertibilidad del producto de matrices.

En efecto, supongamos que la matriz B no es invertible; por el Teorema 7, existe $\mathbf{x} \neq \mathbf{0}$ tal que $B\mathbf{x} = \mathbf{0}$, por lo tanto, existe $\mathbf{x} \neq \mathbf{0}$ tal que $AB\mathbf{x} = \mathbf{0}$, y nuevamente, por el Teorema 7, podemos concluir que AB no es invertible.

De otro lado, si la matriz B es invertible y la matriz A no, entonces existe $\mathbf{y} \neq \mathbf{0}$ tal que $A\mathbf{y} = \mathbf{0}$. Sea $\mathbf{x} = B^{-1}\mathbf{y}$; como $\mathbf{y} \neq \mathbf{0}$, entonces $\mathbf{x} \neq \mathbf{0}$. Además, $AB\mathbf{x} = AB(B^{-1}\mathbf{y}) = A\mathbf{y} = \mathbf{0}$ y por el Teorema 7, podemos concluir que AB no es invertible. Asi, hemos demostrado el siguiente corolario.

Corolario 7.1
Si la matriz AB es invertible y las matrices A y B son cuadradas, entonces las matrices A y B son invertibles.

3.6. Transposición de Matrices

Hasta aquí, las operaciones entre matrices son similares a las de los números reales. Introduzcamos una nueva operación que no tiene análoga en los números reales y se aplica a una sola matriz y que es de mucha utilidad tanto teórica como práctica.

Definición 6 [*Transpuesta de una matriz*]. La transpuesta de una matriz A, de tamaño $m \times n$, es la matriz A^T, de tamaño $n \times m$, que se obtiene tomando la i-ésima columna de A^T como la i-ésima fila de A; es decir, si $A = (a_{ij})$, entonces $A^T = (a_{ji})$.

Notemos que si las columnas de A^T son las filas de A, las filas de A^T son las columnas de A.

Ejemplo 16. Encontremos las transpuestas de las siguientes matrices

$$A = \begin{pmatrix} 1 & 3 & -1 \\ -2 & 0 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -2 & 3 \\ 3 & 0 & -1 \\ -1 & 5 & 7 \end{pmatrix}, \quad C = \begin{pmatrix} 2 \\ 3 \\ -5 \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 3 & -1 & 0 \end{pmatrix}.$$

Al transponer, obtenemos

$$A^T = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ -1 & 5 \end{pmatrix}, \quad B^T = \begin{pmatrix} 1 & 3 & -1 \\ -2 & 0 & 5 \\ 3 & -1 & 7 \end{pmatrix}, \quad C^T = \begin{pmatrix} 2 & 3 & -5 \end{pmatrix} \quad \text{y} \quad D^T = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}.$$

Observemos que, efectivamente, la segunda fila de A, $\begin{pmatrix} -2 & 0 & 5 \end{pmatrix}$, es la segunda columna de A^T, $\begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix}$, y que la segunda columna de A, $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$, es la segunda fila de A^T, $\begin{pmatrix} 3 & 0 \end{pmatrix}$. □

Notemos también que, cuando los vectores son vistos como matrices, la transposición nos permite otra forma de interpretar o de expresar el producto escalar entre dos vectores. Sean

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \quad \text{y} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

vectores de R^n, entonces

$$\begin{aligned} \mathbf{u}\cdot\mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \\ &= \mathbf{u}^T\mathbf{v}. \end{aligned}$$

En forma similar, la transposición de matrices nos permite ver que $\mathbf{e}_i^T A$ es la i-ésima fila de A y que $\mathbf{e}_i^T A\mathbf{e}_j = a_{ij}$, la (i,j)-componente de la matriz A. Su demostración la dejamos como ejercicio para el lector.

Veamos ahora, cuales son las relaciones entre las operaciones básicas y la transposición de matrices.

Teorema 8 [*Propiedades algebraicas de la transpuesta de matrices*].
Sean A y B matrices tales que las operaciones indicadas están bien definidas y λ un número real (escalar). Entonces,

1. $(A^T)^T = A$.
2. $(A+B)^T = A^T + B^T$.
3. $(\lambda A)^T = \lambda A^T$.
4. $(AB)^T = B^TA^T$.(Notemos el orden del producto de las transpuestas)
5. Si A es invertible, A^T también es invertible y $(A^T)^{-1} = (A^{-1})^T$.

Demostración: La demostración de las tres primeras propiedades es sencilla y directa, las cuales dejamos como ejercicio para el lector. Demostremos las dos últimas.

4. Para demostrar esta propiedad, en primer lugar, observemos que para que el producto AB esté bien definido, si A es una matriz $m\times n$, B debe ser una matriz $n\times p$ y entonces AB es una matriz $m\times p$. Así, B^T es una matriz $p\times n$ y A^T es una matriz $n\times m$, de tal forma que el producto B^TA^T está bien definido y es una matriz $p\times m$, al igual que $(AB)^T$. Denotemos por $fil_i(U)$ a la i-ésima fila de U, por $col_j(U)$ a la j-ésima columna de U y por $(U)_{ij}$ a la componente (i,j) de U. Así, con esta notación, tenemos

$$\begin{aligned} ((AB)^T)_{ij} &= (AB)_{ji} \\ &= fil_j(A)\cdot col_i(B) \\ &= col_j(A^T)\cdot fil_i(B^T) \\ &= fil_i(B^T)\cdot col_j(A^T) \\ &= (B^TA^T)_{ij}. \end{aligned}$$

En otras palabras, tenemos que la componente (i,j) de $(AB)^T$ es la componente (i,j) de B^TA^T. Y, debido a que esto se tiene para cualquier i y j , podemos concluir que $(AB)^T = B^TA^T$.

5. Puesto que A es invertible, existe A^{-1}, tal que $AA^{-1} = I$. Ahora, por la Propiedad 4,

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I.$$

Así, que A^T es invertible y su inversa es $(A^T)^{-1} = (A^{-1})^T$. □

Las matrices que son iguales a sus transpuestas, juegan un papel fundamental en la teoría del álgebra lineal, como veremos en los siguientes capítulos. Ellas reciben un nombre especial derivado de la estructura de sus componentes.

Definición 7 [*Matriz simétrica*]. Una Matriz A de tamaño $n \times n$ es simétrica, si y solo si, es igual a su transpuesta; es decir, si y solo si,

$$A = A^T.$$

Ejemplo 17. Determinemos cuáles de las siguientes matrices son simétricas

$$A = \begin{pmatrix} 1 & -3 \\ -3 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & -7 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 4 \\ -1 & 1 & -2 \\ -4 & 2 & 5 \end{pmatrix} \quad y \quad D = \begin{pmatrix} 0 & -1 & 4 & -7 \\ -1 & 1 & -2 & 0 \\ -4 & -2 & 5 & 0{,}5 \end{pmatrix}.$$

Es fácil ver que $A = A^T$ y $B = B^T$, mientras que $C^T \neq C$, por tanto A y B son simétricas, mientras que C no lo es. La matriz D, por no ser cuadrada, tampoco es simétrica. □

Ejemplo 18. Veamos que si A es una matriz de tamaño $m \times n$, AA^T es una matriz simétrica.

Por las Propiedades 1 y 4 del Teorema 8, tenemos que $(AA^T)^T = (A^T)^T A^T = AA^T$; por lo tanto, AA^T es simétrica. □

3.7. Matrices Elementales

Recordemos que el método de eliminación de Gauss consiste en aplicar operaciones elementales entre filas a la matriz aumentada del sistema de ecuaciones lineales hasta obtener una forma escalonada de ella. Veremos que aplicar una operación elemental a una matriz A es equivalente a pre-multiplicar esta matriz por otra, llamada *matriz elemental*.

Definición 8 [*Matriz elemental*]. Llamamos *matriz elemental* de tamaño $n \times n$ a una matriz que se obtiene de aplicar una operación elemental entre filas a la matriz identidad I, también de tamaño $n \times n$.

Como existen tres tipos de operaciones elementales entre filas (*eliminación, escalamiento* y *permutación*), existen tres tipos de matriz elemental que reciben los mismos nombres. La estructura de cada tipo de matriz elemental es muy especial como lo podemos ver en el siguiente ejemplo.

Ejemplo 19. Veamos que las siguientes matrices son elementales.

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

La matriz E_1 se obtiene de la matriz I_3, al aplicarle la operación elemental de Tipo Escalamiento $\mathbf{5}F_{\mathbf{2}} \longrightarrow F_{\mathbf{2}}$, por lo cual E_1 es una matriz elemental de Tipo Escalamiento. Notemos que la única componente de E_1 diferente a las de I_3 está en la posición $(\mathbf{2{,}2})$, que es **5** en lugar de 1, y que esta posición está relacionada con la fila **2** usada en la operación elemental aplicada, y que **5** es el escalar involucrado en dicha operación.

La matriz E_2 se obtiene de la matriz I_2, al aplicarle la operación elemental de Tipo Eliminación $F_{\mathbf{2}} + (\mathbf{-3})F_{\mathbf{1}} \longrightarrow F_{\mathbf{2}}$, por lo cual E_2 es una matriz elemental de Tipo Eliminación. Notemos que la única componente de E_2 diferente a las de I_2 está en la posición $(\mathbf{2{,}1})$, que es **-3** en lugar de 0, y que esta posición está relacionada con las filas **1** y **2** usadas en la operación elemental aplicada, y que **-3** es el escalar involucrado en dicha operación.

La matriz E_3 se obtiene de la matriz I_4, al aplicarle la operación elemental de Tipo Permutación $F_{\mathbf{1}} \longleftrightarrow F_{\mathbf{3}}$, por lo cual E_3 es una matriz elemental de Tipo Permutación. Notemos que las únicas componentes de E_3

diferente a las de I_4 están en las posiciones (**1,1**), (**1,3**), (**3,1**) y (**3,3**) , que donde había un cero quedó un 1 y viceversa, y que estas posiciones están relacionadas con las filas **1** y **3** usadas en la operación elemental aplicada. □

En general, la matriz elemental (Tipo Eliminación) correspondiente a la operación elemental $F_{\mathbf{i}}+(c)F_{\mathbf{j}} \longrightarrow F_{\mathbf{i}}$ difiere de la matriz idéntica sólo en la posición $(\mathbf{i},\mathbf{j})$, donde aparece c en lugar de un cero. Igualmente, la matriz elemental (Tipo Escalamiento) correspondiente a la operación elemental $(c)F_{\mathbf{i}} \longrightarrow F_{\mathbf{i}}$ difiere de la matriz idéntica sólo en la posición $(\mathbf{i},\mathbf{i})$, donde aparece c en lugar de uno. Y, de la misma forma, la matriz elemental (Tipo Permutación) correspondiente a la operación elemental $F_{\mathbf{i}} \longleftrightarrow F_{\mathbf{j}}$ difiere de la matriz idéntica en las posiciones $(\mathbf{i},\mathbf{i})$, $(\mathbf{i},\mathbf{j}), (\mathbf{j},\mathbf{i})$ y $(\mathbf{j},\mathbf{j})$, donde aparecen unos en lugar de ceros y viceversa.

Ahora, observemos que pre-multiplicar por una de estas matrices elementales es equivalente a aplicar la operación elemental correspondiente.

Ejemplo 20. Sean las matrices

$$A = \begin{pmatrix} 0 & 0 & 2 \\ 3 & -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 2 & 5 \end{pmatrix} \quad \text{y} \quad C = \begin{pmatrix} 3/4 & 0 & -1/2 \\ -1 & -1 & 3 \\ 1 & 2 & -1 \\ 4 & 0 & 5 \end{pmatrix}.$$

Al aplicar la operación elemental entre filas $F_1 \longleftrightarrow F_2$ a A, obtenemos

$$A_1 = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

y, al pre-multiplicar a A por E_1, la correspondiente matriz elemental Tipo Permutación, obtenemos

$$E_1A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 2 \\ 3 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = A_1.$$

De igual forma, al aplicar la operación elemental entre filas $F_3 + 2F_2 \longrightarrow F_3$ a B, obtenemos

$$B_1 = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 7 \end{pmatrix},$$

y, al pre-multiplicar a B por E_2, la correspondiente matriz elemental Tipo Eliminación, obtenemos

$$E_2B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 7 \end{pmatrix} = B_1.$$

Y similarmente, al aplicar la operación elemental entre filas $4F_1 \longrightarrow F_1$ a C, obtenemos

$$C_1 = \begin{pmatrix} 3 & 0 & -2 \\ -1 & -1 & 3 \\ 1 & 2 & -1 \\ 4 & 0 & 5 \end{pmatrix},$$

y, al pre-multiplicar a C por E_3, la correspondiente matriz elemental Tipo Escalamiento, obtenemos

$$E_3C = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3/4 & 0 & -1/2 \\ -1 & -1 & 3 \\ 1 & 2 & -1 \\ 4 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 0 & -2 \\ -1 & -1 & 3 \\ 1 & 2 & -1 \\ 4 & 0 & 5 \end{pmatrix} = C_1.$$

□

Observemos que el orden o tamaño de las matrices elementales es $m \times m$, donde m es el número de filas de la matriz a la cual se le aplica la correspondiente operación elemental entre filas. Observemos también que las operaciones elementales entre filas son reversibles; es decir, que es posible recuperar la matriz a la cual se le ha aplicado una operación elemental entre filas, aplicando al resultado otra operación elemental entre filas. En efecto, si A_1 se obtiene al aplicarle la operación elemental $F_i + cF_j \longrightarrow F_i$ a A, entonces A se obtiene al aplicarle la operación elemental $F_i - cF_j \longrightarrow F_i$ a A_1. De igual forma, si B_1 se obtiene de aplicarle la operación elemental $cF_i \longrightarrow F_i$ a B, entonces B se obtiene de aplicarle la operación elemental $\frac{1}{c}F_i \longrightarrow F_i$ a B_1 y, finalmente, si C_1 se obtiene de aplicarle la operación elemental $F_i \longleftrightarrow F_j$ a C, entonces C se obtiene de aplicarle la misma operación elemental $F_i \longleftrightarrow F_j$ a C_1. Esto nos indica que las matrices elementales son invertibles y, aún más, que sus inversas son matrices elementales del mismo tipo, como se resume en el siguiente teorema.

Teorema 9 [*Inversa de las matrices elementales*].
Las matrices elementales son invertibles y la inversa de una matriz elemental es a su vez una matriz elemental del mismo tipo; más exactamente,

1. La inversa de la matriz elemental correspondiente a la operación elemental $F_i + cF_j \longrightarrow F_i$ es la matriz elemental correspondiente a la operación elemental $F_i - cF_j \longrightarrow F_i$.

2. La inversa de la matriz elemental correspondiente a la operación elemental $cF_i \longrightarrow F_i,\ c \neq 0$ es la matriz elemental correspondiente a la operación elemental $\dfrac{1}{c}F_i \longrightarrow F_i$.

3. La inversa de la matriz elemental correspondiente a la operación elemental $F_i \longleftrightarrow F_j$ es la matriz elemental correspondiente a la misma operación elemental $F_i \longleftrightarrow F_j$.

Demostración: En cada uno de los tres casos, la demostración se hace por cálculo directo. □

Ejemplo 21. Veamos cuales son las inversas de cada una de las siguientes matrices elementales

$$E_1 = \begin{pmatrix} 1 & 0 \\ \mathbf{-3} & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 & \mathbf{1} & 0 \\ 0 & 1 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Por el resultado anterior y teniendo en cuenta el tipo de matriz elemental de cada una (eliminación, escalamiento y permutación, respectivamente), tenemos que

$$E_1^{-1} = \begin{pmatrix} 1 & 0 \\ \mathbf{3} & 1 \end{pmatrix}, \quad E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{1/5} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_3^{-1} = \begin{pmatrix} 0 & 0 & \mathbf{1} & 0 \\ 0 & 1 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = E_3.$$

□

Volviendo al método de eliminación de Gauss, es claro que siempre es posible obtener una forma escalonada de una matriz dada con operaciones elementales solamente de Tipo Eliminación y Permutación. Así, tenemos que, a partir de una matriz dada A, es posible obtener una matriz triangular superior (o con forma escalonada) U, pre-multiplicando sucesivamente por matrices elementales apropiadas de Tipo Eliminación y Permutación a la matriz A.

Ejemplo 22. (Parte A). Dada la matriz

$$A = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{pmatrix},$$

calculemos una matriz U de forma escalonada equivalente a A, y las matrices elementales $E_1, \cdots, E_k$, tales que

$$E_k \cdots E_1 A = U.$$

Al aplicar operaciones elementales, tenemos que

$$\begin{pmatrix} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{pmatrix}$$

$$F_2 - F_1 \longrightarrow F_2 \quad \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \\ 0 & 2 & -2 & 1 \end{pmatrix}$$

$$F_3 - \tfrac{2}{3}F_2 \longrightarrow F_3 \quad \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Así que

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{pmatrix}}_{E_2} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{E_1} \underbrace{\begin{pmatrix} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \\ 0 & 0 & 0 & 5 \end{pmatrix}}_{U}.$$

□

Si, a la matriz obtenida en el proceso de Eliminación de Gauss, continuamos aplicándole las operaciones elementales (Tipo Escalamiento y Tipo Eliminación) requeridas en el Método de Sustitución hacia atrás, obtenemos una matriz de forma escalonada donde las columnas pivote son $\mathbf{e}_1, \ \mathbf{e}_2, \ldots, \ \mathbf{e}_k$, los primeros vectores canónicos de R^n, como pudimos apreciar en el capítulo anterior e ilustramos en la parte B del Ejemplo 22 que presentamos a continuación[4].

Ejemplo 22. (Parte B). A la matriz obtenida en la Parte A del Ejemplo 22, apliquémosle las operaciones elementales del Método de Sustitución hacia atrás para obtener su forma escalonada donde las columnas pivote son los primeros vectores canónicos de R^n *(matriz escalonada reducida).*

$$\begin{matrix} \frac{1}{5}F_3 & \longrightarrow & F_3 \\ F_2 + 6F_3 & \longrightarrow & F_2 \\ F_1 - 3F_3 & \longrightarrow & F_1 \end{matrix} \quad \begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{matrix} \frac{1}{3}F_2 & \longrightarrow & F_2 \\ F_1 + F_2 & \longrightarrow & F_1 \end{matrix} \quad \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Si llamamos $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{5} \end{pmatrix}$, $E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{pmatrix}$, $E_5 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $E_7 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ las matrices elementales correspondientes a las anteriores operaciones elementales,

[4]Muchos autores hacen un uso excesivo de este tipo especial de forma escalonada y la llaman *forma escalonada reducida.*

podemos resumir la parte A y B de este ejemplo con la siguiente expresión

$$(E_7\ E_6\ E_5\ E_4)\ (E_3\ E_2\ E_1\ A) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Observemos que las columnas pivotales en esta matriz (primera, segunda y cuarta) son precisamente $\mathbf{e}_1, \mathbf{e}_2$ y $\mathbf{e}_3$, los vectores canónicos de R^3, en su orden. □

Cuando A es una matriz invertible de tamaño $n \times n$, cualquier matriz escalonada equivalente tiene n pivotes (Teorema 6); por lo tanto, al continuar aplicando las operaciones elementales del Método de Sustitución hacia atrás, obtendremos como columnas pivote $\mathbf{e}_1,\ \mathbf{e}_2, \ldots,\ \mathbf{e}_n$, los n vectores canónicos de R^n; es decir, obtendremos la matriz idéntica de tamaño $n \times n$. Esta importante propiedad de las matrices invertibles la resumimos en el siguiente teorema.

Teorema 10 [*Caracterización de la inversa en término de matrices elementales*].
Una matriz A de tamaño $n \times n$ es invertible, si y solo si, la matriz A es el producto de matrices elementales.

Demostración: Si la matriz A es invertible, al aplicar sucesivamente los algoritmos de Eliminación de Gauss y Sustitución hacia atrás, como se describió previamente, obtendremos la matriz idéntica; es decir, existen matrices elementales $E_1,\ E_2,\ \ldots,\ E_k$, donde $k \leq n^2$, tales que

$$E_k \cdots E_2 E_1 A = I.$$

Pre-multiplicando por las matrices elementales inversas (en el orden apropiado) ambos lados de la expresión anterior, obtenemos

$$A = E_1^{-1} E_2^{-1} \cdots E_k^{-1} I,$$

lo cual está expresando a la matriz A como el producto de matrices elementales, ya que la matriz inversa de una matriz elemental es también una matriz elemental.

En el otro sentido, es obvio que si la matriz A es el producto de matrices elementales, la matriz A es invertible, puesto que las matrices elementales son invertibles y, por el Teorema 6, el producto de matrices invertibles es invertible. □

3.8. Factorización LU

Así como los números se pueden factorizar y de esta factorización se pueden determinar algunas propiedades de ellos, la factorización de matrices es de mucha utilidad tanto práctica como teórica. La que veremos en seguida es la llamada factorización LU, la cual se obtiene de una formulación matricial del algoritmo de eliminación de Gauss y nos proporciona una forma más eficiente de calcular la solución de un sistema de ecuaciones lineales.

En la Eliminación de Gauss, cuando no son necesarias las operaciones o matrices elementales de Tipo Permutación, como en el Ejemplo 22 (Parte A), obtenemos otro resultado importante[5]. En efecto, al pre-multiplicar por las matrices elementales inversas (en el orden apropiado) ambos lados de la expresión

$$E_k \cdots E_1 A = U,$$

obtenemos

$$A = E_1^{-1} \cdots E_k^{-1} U = LU,$$

[5]Cuando, en la Eliminación de Gauss, se utilizan operaciones o matrices elementales de Tipo Permutación, se obtiene un resultado similar al descrito en esta parte, el cual se conoce con el nombre de *Factorización PLU* de A, donde P es una matriz de permutación, L es una matriz triangular inferior con unos en la diagonal y U es una matriz triangular superior, tales que $PA = LU$; es decir, obtenemos la factorización LU de la matriz que se obtiene de A permutando adecuadamente sus filas. Una descripción detallada de esta factorización aparece en el Apéndice como un ejemplo adicional de las aplicaciones del álgebra matricial.

donde U es una matriz triangular superior y $L = E_1^{-1} \cdots E_k^{-1}$ es una matriz triangular inferior, cuadrada e invertible, con unos en la diagonal. Además, los elementos diferentes de cero de la matriz L son los opuestos de los escalares o multiplicadores involucrados en las operaciones elementales Tipo Eliminación usadas en la eliminación de Gauss.

En efecto, obtener la matriz L **no** requiere cálculo adicional como parece hacer ver la fórmula que la define. La matriz L es muy fácil de obtener gracias a que el producto de matrices elementales del Tipo Eliminación es prácticamente una *superposición de escalares* sobre una matriz idéntica cuando el factor izquierdo tiene su escalar en la misma columna o en una columna que antecede a la columna donde está el escalar del factor derecho. Observemos que las matrices elementales que definen a L tienen esta propiedad (el opuesto de los escalares o multiplicadores aparece porque la matriz L está definida como producto de las inversas de las matrices elementales). Ilustremos estos hechos con un ejemplo.

Ejemplo 23. Sean E_1, E_2 y E_3 las matrices elementales de orden 4×4 correspondientes a las operaciones elementales $F_3 + 2F_2 \longrightarrow F_3$, $F_4 + (-3)F_2 \longrightarrow F_4$ y $F_4 + 5F_3 \longrightarrow F_4$, respectivamente. Observemos la *superposición*, sobre una matriz idéntica, de los escalares ($2,\ -3$ y 5) involucrados en las operaciones elementales, al efectuar los productos E_1E_2, E_2E_3 y $E_1E_2E_3$, y que no se tiene en el producto E_3E_1.

$$E_1E_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{-3} & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{2} & 1 & 0 \\ 0 & \mathbf{-3} & 0 & 1 \end{pmatrix}.$$

$$E_2E_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{-3} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \mathbf{5} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{-3} & \mathbf{5} & 1 \end{pmatrix}.$$

$$E_1E_2E_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{-3} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \mathbf{5} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{2} & 1 & 0 \\ 0 & \mathbf{-3} & \mathbf{5} & 1 \end{pmatrix}.$$

$$E_3E_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \mathbf{5} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{2} & 1 & 0 \\ 0 & \mathbf{10} & 0 & 1 \end{pmatrix}.$$

Si estas tres operaciones elementales fuesen las requeridas para escalonar una matriz dada, la matriz L mencionada anteriormente sería

$$L = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{-2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{3} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \mathbf{-5} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \mathbf{-2} & 1 & 0 \\ 0 & \mathbf{3} & \mathbf{-5} & 1 \end{pmatrix}.$$

□

Todo lo anterior lo podemos formalizar en el siguiente teorema.

Teorema 11 [*Factorización o descomposición LU*].
Sea A una matriz de tamaño $m \times n$ que puede reducirse a la matriz escalonada U usando sólo operaciones elementales de Tipo Eliminación. Entonces, podemos factorizar a A como

$$A = LU,$$

donde U es una matriz escalonada $m \times n$ equivalente a A y L es una matriz triangular inferior con unos en la diagonal de tamaño $m \times m$, cuya componente $l_{ij} = -c$ para $i > j$, siendo c el escalar o *multiplicador*

involucrado en la operación elemental Tipo Eliminación $F_i + cF_j \longrightarrow F_i$, que utilizamos para escalonar la matriz A. A esta factorización se le llama *factorización LU* o *descomposición LU*. $\square$

Entre otros usos, esta factorización permite resolver el sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$, reduciendo su solución a la solución de dos sistemas de ecuaciones lineales que podemos resolver fácilmente. Sabiendo que $A = LU$, el sistema $A\mathbf{x} = \mathbf{b}$ se convierte en $LU\mathbf{x} = \mathbf{b}$, lo cual es equivalente a $L\mathbf{y} = \mathbf{b}$, donde $U\mathbf{x} = \mathbf{y}$; por lo tanto, para calcular $\mathbf{x}$, primero resolvemos

$$L\mathbf{y} = \mathbf{b}$$

y a continuación resolvemos

$$U\mathbf{x} = \mathbf{y}.$$

Observemos que los dos sistemas de ecuaciones lineales anteriores son fáciles de resolver. Por ser L una matriz triangular inferior, el primer sistema se resuelve mediante *sustitución hacia adelante*; además, por ser cuadrada y tener unos en la diagonal, es una matriz invertible; por lo tanto, su solución siempre es única. Por ser U una matriz triangular superior, el segundo sistema, se resuelve mediante *sustitución hacía atrás*. Veamos un ejemplo.

Ejemplo 24. Sea $A = \begin{pmatrix} 6 & -3 & 0 & 3 \\ -6 & 3 & 8 & -3 \\ 4 & -2 & -6 & 3 \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} -8 \\ 8 \\ -5 \end{pmatrix}$. Encontremos la factorización LU de A y luego calculemos las soluciones de $A\mathbf{x} = \mathbf{b}$.

Aplicando sucesivamente las operaciones elementales entre filas a la matriz A

$$\begin{array}{c} F_2 + F_1 \longrightarrow F_2 \\ F_3 - \frac{2}{3}F_1 \longrightarrow F_3 \\ F_3 + \frac{6}{8}F_2 \longrightarrow F_3 \end{array} \quad \text{obtenemos} \quad \begin{pmatrix} 6 & -3 & 0 & 3 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = U.$$

Las matrices elementales utilizadas, en su orden, son

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2/3 & 0 & 1 \end{pmatrix} \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3/4 & 1 \end{pmatrix};$$

por lo tanto,

$$L = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2/3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3/4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2/3 & -3/4 & 1 \end{pmatrix},$$

la cual, en la práctica, obtenemos directamente de los multiplicadores o escalares (1, -2/3 y 3/4) usados en las operaciones elementales, como se definió en el Teorema 11.

Ahora, utilicemos la factorización LU de A para resolver el sistema dado. Primero, resolvemos el sistema $L\mathbf{y} = \mathbf{b}$ mediante sustitución hacia adelante, lo cual es equivalente a aplicar las operaciones elementales

$$\begin{array}{c} F_2 + F_1 \longrightarrow F_2 \\ F_3 - \frac{2}{3}F_1 \longrightarrow F_3 \\ F_3 + \frac{3}{4}F_2 \longrightarrow F_3 \end{array} \quad \text{a la matriz} \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & -8 \\ -1 & 1 & 0 & 8 \\ 2/3 & -3/4 & 1 & -5 \end{array}\right), \quad \text{obteniendo} \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & -8 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/3 \end{array}\right);$$

es decir que $\mathbf{y} = (-8 \;\; 0 \;\; 1/3)^T$. Continuamos resolviendo el sistema $U\mathbf{x} = \mathbf{y}$ mediante sustitución hacia atrás, lo cual es equivalente a aplicar las operaciones elementales

$$\begin{array}{c} F_1 - 3F_3 \longrightarrow F_1 \\ \frac{1}{8}F_2 \longrightarrow F_2 \\ \frac{1}{6}F_1 \longrightarrow F_1 \end{array} \quad \text{a la matriz} \quad \left(\begin{array}{cccc|c} 6 & -3 & 0 & 3 & -8 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1/3 \end{array}\right), \quad \text{obteniendo} \quad \left(\begin{array}{cccc|c} 1 & -1/2 & 0 & 0 & -3/2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1/3 \end{array}\right);$$

de donde, concluimos que

$$\mathbf{x} = \begin{pmatrix} -\frac{3}{2} + \frac{1}{2}t \\ t \\ 0 \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ 0 \\ 0 \\ \frac{1}{3} \end{pmatrix} + t \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad t \in R.$$

Supongamos que, posteriormente a lo anterior, necesitamos calcular las soluciones de $A\mathbf{x} = \mathbf{c}$, para $\mathbf{c} = (-6\ \ -2\ \ 2)^T$. Como ya tenemos la factorización LU de A, no necesitamos empezar de nuevo; solo necesitamos resolver los sistemas $L\mathbf{y} = \mathbf{c}$ y $U\mathbf{x} = \mathbf{y}$, como en el caso anterior. Esto es, resolvemos el sistema $L\mathbf{y} = \mathbf{c}$ mediante sustitución hacia adelante, lo cual es equivalente a aplicar las operaciones elementales

$$\begin{array}{c} F_2 + F_1 \longrightarrow F_2 \\ F_3 - \frac{2}{3}F_1 \longrightarrow F_3 \\ F_3 + \frac{3}{4}F_2 \longrightarrow F_3 \end{array} \quad \text{a la matriz} \quad \left(\begin{array}{rrr|r} 1 & 0 & 0 & -6 \\ -1 & 1 & 0 & -2 \\ 2/3 & -3/4 & 1 & 2 \end{array}\right), \quad \text{obteniendo} \quad \left(\begin{array}{rrr|r} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & -8 \\ 0 & 0 & 1 & 0 \end{array}\right);$$

es decir que $\mathbf{y} = (-6\ \ -8\ \ 0)^T$. Continuamos resolviendo el sistema $U\mathbf{x} = \mathbf{y}$ mediante sustitución hacia atrás, lo cual es equivalente a aplicar las operaciones elementales

$$\begin{array}{c} F_1 - 3F_3 \longrightarrow F_1 \\ \frac{1}{8}F_2 \longrightarrow F_2 \\ \frac{1}{6}F_1 \longrightarrow F_1 \end{array} \quad \text{a la matriz} \quad \left(\begin{array}{rrrr|r} 6 & -3 & 0 & 3 & -6 \\ 0 & 0 & 8 & 0 & -8 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right), \quad \text{obteniendo} \quad \left(\begin{array}{rrrr|r} 1 & -1/2 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right);$$

de donde, concluimos que

$$\mathbf{x} = \begin{pmatrix} -1 + \frac{1}{2}t \\ t \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} \frac{1}{2} \\ 1 \\ -0 \\ 0 \end{pmatrix}, \quad t \in R.$$

□

Para poder factorizar una matriz A, como vimos en el ejemplo anterior, es necesario conocer las operaciones elementales necesarias para encontrar una forma escalonada equivalente a ella, lo que usualmente se realiza cuando resolvemos un sistema de ecuaciones lineales con matriz de coeficientes A. Además, resolver el sistema $L\mathbf{y} = \mathbf{b}$ requiere tantas operaciones como las requeridas al incluir el vector $\mathbf{b}$ en el proceso de reducción de A a su forma escalonada, como lo hicimos en el Capítulo 1 (proceso de Eliminación de Gauss aplicado a la matriz aumentada del sistema).

En otras palabras, es claro que la ventaja de la factorización no la apreciamos cuando se resuelve un sistema $A\mathbf{x} = \mathbf{b}$ o varios sistemas simultáneamente, sino cuando necesitamos resolver otro sistema de ecuaciones lineales de la forma $A\mathbf{x} = \mathbf{c}$, después de haber resuelto el sistema $A\mathbf{x} = \mathbf{b}$ o cuando, simultáneamente, necesitamos resolver los dos sistemas, pero el vector de términos independientes de uno de los sistemas depende de la solución del otro, como ocurre en el Método Simplex para resolver problemas de programación lineal, en la estimación del Número Condicional de una matriz, en el proceso de Refinamiento Iterativo de solución de sistemas de ecuaciones lineales y en muchas otras aplicaciones prácticas. En resumen, la factorización LU es el procedimiento alternativo para ahorrar tiempo de computo cuando necesitamos resolver dos o más sistemas de ecuaciones lineales y la solución simultánea de ellos no es posible, como ya lo habíamos planteado al final del Capítulo 1.

Ejemplo 25. Encontremos dos vectores $\mathbf{x}$ y $\mathbf{z}$ de R^4 tales que $A\mathbf{x} = \mathbf{b}$ y $\mathbf{z}^T A = \mathbf{c}^T$, donde

$$A = \begin{pmatrix} 2 & -1 & 3 & 0 \\ -2 & 4 & -3 & 5 \\ 0 & 9 & -1 & 18 \\ 4 & -5 & 1 & 8 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -5 \\ 10 \\ 21 \\ 13 \end{pmatrix} \quad \text{y} \quad \mathbf{c} = \begin{pmatrix} 2 \\ -4 \\ 1 \\ 5 \end{pmatrix}.$$

Observemos que para calcular $\mathbf{z}$, debemos resolver el sistema de ecuaciones lineales $A^T\mathbf{z} = \mathbf{c}$; por lo tanto, no es posible calcular a $\mathbf{x}$ y a $\mathbf{z}$ simultáneamente.

Primero, resolvamos los dos sistemas, $A\mathbf{x} = \mathbf{b}$ y $A^T\mathbf{z} = \mathbf{c}$, utilizando, para cada uno de ellos, el algoritmo de eliminación de Gauss seguido del algoritmo de sustitución hacia atrás y luego los resolvemos utilizando la factorización LU de la matriz A.

Para resolver el primer sistema, aplicamos sucesivamente las operaciones elementales $F_2 + F_1 \longrightarrow F_2$, $F_4 - 2F_1 \longrightarrow F_4$, $F_3 - 3F_2 \longrightarrow F_3$, $F_4 + F_2 \longrightarrow F_4$ y $F_4 - 5F_3 \longrightarrow F_4$ a la matriz aumentada $[A|\mathbf{b}]$ y obtenemos la matriz escalonada

$$[U|\mathbf{y}] = \left(\begin{array}{rrrr|r} 2 & -1 & 3 & 0 & -5 \\ 0 & 3 & 0 & 5 & 5 \\ 0 & 0 & -1 & 3 & 6 \\ 0 & 0 & 0 & -2 & -2 \end{array}\right).$$

Finalmente, aplicando el algoritmo de sustitución hacia atrás, tenemos que $\mathbf{x} = (2 \;\; 0 \;\; -3 \;\; 1)^T$.

Para resolver el segundo sistema, aplicamos sucesivamente las operaciones elementales $F_2 + 0{,}5F_1 \longrightarrow F_2$, $F_3 - 1{,}5F_1 \longrightarrow F_3$, $F_4 - \frac{5}{3}F_2 \longrightarrow F_4$ y $F_4 + 3F_3 \longrightarrow F_4$ a la matriz aumentada $[A^T|\mathbf{c}]$ y obtenemos la matriz escalonada

$$[V|\mathbf{r}] = \left(\begin{array}{rrrr|r} 2 & -2 & 0 & 4 & 2 \\ 0 & 3 & 9 & -3 & -3 \\ 0 & 0 & -1 & -5 & -2 \\ 0 & 0 & 0 & -2 & 4 \end{array}\right).$$

Finalmente, aplicando el algoritmo de sustitución hacia atrás, tenemos que $\mathbf{z} = (-30 \;\; 39 \;\; 12 \;\; -2)^T$.

Ahora, resolvamos el mismo problema utilizando la factorización LU de A, para lo cual, debemos encontrar una forma escalonada equivalente a la matriz A, utilizando el algoritmo de eliminación de Gauss. Aplicando sucesivamente las operaciones elementales $F_2+F_1 \longrightarrow F_2$, $F_4-2F_1 \longrightarrow F_4$, $F_3-3F_2 \longrightarrow F_3$, $F_4+F_2 \longrightarrow F_4$ y $F_4 - 5F_3 \longrightarrow F_4$ a la matriz A, obtenemos la matriz escalonada

$$U = \begin{pmatrix} 2 & -1 & 3 & 0 \\ 0 & 3 & 0 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Teniendo en cuenta las operaciones elementales utilizadas y la definición de la matriz L, tenemos que

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 2 & -1 & 5 & 1 \end{pmatrix}.$$

Como $A = LU$, para calcular $\mathbf{x}$, resolvemos, mediante sustitución hacia adelante, el sistema $L\mathbf{y} = \mathbf{b}$, de donde obtenemos que $\mathbf{y} = (-5 \;\; 5 \;\; 6 \;\; -2)^T$ y, mediante sustitución hacia atrás, el sistema $U\mathbf{x} = \mathbf{y}$, para finalmente obtener que $\mathbf{x} = (2 \;\; 0 \;\; -3 \;\; 1)^T$.

Como $A^T = U^TL^T$, para calcular z, resolvemos, mediante sustitución hacia adelante, el sistema $U^T\mathbf{w} = \mathbf{c}$, de donde obtenemos que $\mathbf{w} = (1 \;\; -1 \;\; 2 \;\; -2)^T$ y, mediante sustitución hacia atrás, el sistema $L^T\mathbf{z} = \mathbf{w}$, para finalmente obtener que $\mathbf{z} = (-30 \;\; 39 \;\; 12 \;\; -2)^T$. $\square$

De las dos formas presentadas para resolver el problema anterior, observemos que el cálculo del vector $\mathbf{x}$ es prácticamente idéntico en ambas. Efectivamente, en las dos formas, se aplican las mismas operaciones elementales para encontrar la matriz escalonada equivalente a A y se resuelve el mismo sistema triangular $U\mathbf{x} = \mathbf{y}$ usando el algoritmo de sustitución hacia atrás. La diferencia aparente está en que, en la primera forma, las operaciones elementales se aplican a la matriz aumentada para obtener el vector $\mathbf{y}$, mientras que

en la segunda, las operaciones elementales se aplican sólo a la matriz de coeficientes del sistema y luego, se resuelve el sistema triangular $L\mathbf{y} = \mathbf{b}$ usando el algoritmo de sustitución hacia adelante, procedimientos que requieren exactamente las mismas operaciones. La diferencia *(ventaja)* de calcular la factorización LU de la matriz A es que evita tener que utilizar el algoritmo de eliminación de Gauss para el cálculo del vector $\mathbf{z}$, lo cual es importante si tenemos en cuenta que es precisamente este algoritmo la parte que más cálculos requiere en el proceso de resolver un sistema de ecuaciones lineales.

3.9. Determinantes

El concepto de determinante surgió independientemente del de matriz y su teoría fue desarrollada, hace casi dos siglos, para solucionar problemas prácticos. Hoy en día, ha perdido su importancia práctica por su costo en tiempo de cálculo y por el desarrollo de otras técnicas de computo que lo reemplazan. No obstante, por su importancia teórica, presentamos el concepto de determinante y sus propiedades más usadas.

Para empezar, digamos que el determinante de una matriz *(cuadrada)* es un número real, el cual definiremos en término de determinantes de matrices de tamaño menor, es decir, usaremos la recurrencia en la definición de determinante. Antes de introducir esta definición, presentaremos el conceptos de *menor de una matriz.* Posteriormente, como una forma de simplificar la expresión que define un determinante, introduciremos la noción de *cofactor de una matriz.* Este último concepto también será utilizado, al final de esta sección, para definir la *matriz adjunta.*

Definición 9 [*Menor*]. Dada una matriz A de tamaño $n \times n$, definimos M_{ij}, *el menor* (i, j) de A, como la matriz de tamaño $(n-1) \times (n-1)$ que resulta al eliminar la fila i y la columna j de la matriz A.

Ejemplo 26. Encuentre los menores M_{23} y M_{32} de la matriz $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{pmatrix}$.

Al eliminar la segunda fila y tercera columna de A, obtenemos M_{23}, el menor $(2, 3)$ de la matriz A,

$$M_{23} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

y al eliminar la tercera fila y segunda columna de A, obtenemos M_{32}, el menor $(3, 2)$ de la matriz A,

$$M_{32} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}.$$

□

Definición 10 [*Determinante*]. Sea $A = (a_{ij})$ una matriz $n \times n$. Definimos $\det(\alpha)$, *el determinante* de una matriz 1×1, como α para todo $\alpha \in R$ y $\det A$, *el determinante*[6] de A, como la suma de los productos de a_{1j}, la j-ésima componente de la primera fila de la matriz A, por $\det M_{1j}$, el determinante del menor $(1, j)$ de A, multiplicado por $(-1)^{1+j}$; es decir,

$$\begin{array}{rcl} \det(\alpha) & = & \alpha, \qquad \alpha \in R \\ \det A & = & a_{11} \det M_{11} - a_{12} \det M_{12} + \ldots + (-1)^{1+n} a_{1n} \det M_{1n}. \end{array} \tag{3.1}$$

Ejemplo 27. Calculemos el determinante de una matriz 2×2.
Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Entonces, por la definición anterior, tenemos

$$\det A = a \det M_{11} - b \det M_{12} = a \det(d) - b \det(c) = ad - bc.$$

□

[6] También utilizaremos $|A|$ como notación para el determinante de la matriz A

Para simplificar la expresión (3.1), la cual define el determinante de una matriz $n \times n$, introduzcamos la siguiente definición, que, como dijimos en la introducción, también será usada al final de esta sección, para definir la matriz adjunta.

Definición 11 [*Cofactor*]. Dada una matriz A de tamaño $n \times n$, definimos A_{ij}, *el cofactor* (i, j) de A, como el escalar que resulta al multiplicar el determinante del menor (i, j) de A por $(-1)^{i+j}$, es decir,

$$A_{ij} = (-1)^{i+j} \det M_{ij}.$$

Ejemplo 28. Dada la matriz A del Ejemplo 26, calcule los cofactores A_{23} y A_{33}.

Según la definición anterior y usando el resultado del ejemplo anterior,

$$A_{23} = (-1)^{2+3} \det M_{23} = - \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = - [1 \cdot (-1) - 1 \cdot 1] = 2$$

y

$$A_{33} = (-1)^{3+3} \det M_{33} = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = (1 \cdot 3 - 1 \cdot 2) = 1$$

□

Hasta el momento, debemos tener claro que un menor de una matriz $n \times n$ es una matriz $(n-1) \times (n-1)$ y que un cofactor de una matriz es un número real. Además, veamos que, usando la definición de cofactor, la expresión (3.1), para el caso de una matriz $n \times n$, se reduce a la suma de los productos de las componentes de la primera fila de A por sus cofactores correspondientes; es decir,

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n}, \qquad (3.2)$$

lo cual nos será útil, especialmente, en los desarrollos teóricos sobre determinantes.

Ejemplo 29. Calculemos el determinante de cada una de las siguientes matrices:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 1 & -1 & -2 & 0 \\ 7 & 1/2 & 4 & -1 \end{pmatrix}.$$

Usando (3.2) y el resultado del Ejemplo 27, tenemos

$$\begin{aligned} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{vmatrix} &= a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \\ &= 1(-1)^{1+1} \begin{vmatrix} 3 & 1 \\ -1 & -2 \end{vmatrix} + 1(-1)^{1+2} \begin{vmatrix} 2 & 1 \\ 1 & -2 \end{vmatrix} + 1(-1)^{1+3} \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} \\ &= 1\ [3(-2) - 1(-1)] - 1\ [2(-2) - 1 \cdot 1] + 1\ [2(-1) - 3 \cdot 1] \\ &= -5 + 5 - 5 = -5 \end{aligned}$$

$$\begin{aligned} \begin{vmatrix} \mathbf{1} & 0 & 0 & 0 \\ 2 & \mathbf{3} & 0 & 0 \\ 1 & -1 & \mathbf{-2} & 0 \\ 7 & 1/2 & 4 & \mathbf{-1} \end{vmatrix} &= b_{11}B_{11} + b_{12}B_{12} + b_{13}B_{13} + b_{14}B_{14} \\ &= 1 \begin{vmatrix} 3 & 0 & 0 \\ -1 & -2 & 0 \\ 1/2 & 4 & -1 \end{vmatrix} - 0 \begin{vmatrix} 2 & 0 & 0 \\ 1 & -2 & 0 \\ 7 & 4 & -1 \end{vmatrix} + 0 \begin{vmatrix} 2 & 3 & 0 \\ 1 & -1 & 0 \\ 7 & 1/2 & -1 \end{vmatrix} - 0 \begin{vmatrix} 2 & 3 & 0 \\ 1 & -1 & -2 \\ 7 & 1/2 & 4 \end{vmatrix} \\ &= 1 \left\{ 3 \begin{vmatrix} -2 & 0 \\ 4 & -1 \end{vmatrix} - 0 \begin{vmatrix} -1 & 0 \\ 1/2 & -1 \end{vmatrix} + 0 \begin{vmatrix} -1 & -2 \\ 1/2 & 4 \end{vmatrix} \right\} \\ &= 1 \cdot 3\,[(-2)(-1) - 0 \cdot 4] = \mathbf{1} \cdot \mathbf{3(\text{-}2)(\text{-}1)} = 6 \end{aligned}$$

□

Podemos ver que calcular un determinante es más tedioso, en la medida que el tamaño de la matriz sea más grande. Notemos que el cálculo del determinante de una matriz 2×2 requiere 2 multiplicaciones y una suma. El cálculo del determinante de una matriz 3×3 implica el cálculo del determinante de 3 matrices 2×2. El cálculo del determinante de una matriz 4×4 implica el cálculo del determinante de 4 matrices 3×3 y por tanto, el cálculo del determinante de $4\cdot 3 = 12$ matrices 2×2. Así, por ejemplo, el cálculo del determinante de una matriz 10×10 implica el cálculo del determinante de $10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3 = 1'814{,}400$ matrices 2×2, lo cual requiere al menos $2\cdot 1'814{,}400 = 3'628{,}800$ multiplicaciones (además de un número similar de sumas)[7]. Esta enorme cantidad de operaciones nos obliga a buscar propiedades que nos permitan diseñar métodos más eficientes para calcular un determinante.

Un resultado sorprendente es que al calcular la suma del producto de las componentes de cualquier fila o columna de una matriz por sus cofactores correspondientes, obtenemos el mismo resultado, como lo enunciamos en el siguiente teorema, conocido como la Expansión de Laplace y cuya demostración omitimos, ya que es un poco extensa.

Teorema 12 [*Desarrollo o Expansión de Laplace*].
Dada $A = (a_{ij})$ una matriz $n\times n$,

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}$$

para cualquier $i = 1, 2, \ldots, n$, y

$$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}$$

para cualquier $j = 1, 2, \ldots, n$. (estas formas de calcular el determinante las llamaremos desarrollo por cofactores utilizando la i-ésima fila, en el primer caso, y la j-ésima columna, en el segundo). □

Este teorema nos permite concluir que el determinante de una matriz y su transpuesta es el mismo.

Corolario 12.1
Dada A, una matriz $n\times n$,

$$\det A = \det A^T.$$

□

Ejemplo 30. Calculemos el determinante de la matriz $A = \begin{pmatrix} 2 & -1 & 0 & 3 \\ 0 & 4 & -2 & 0 \\ -3 & 0 & 1 & 2 \\ 0 & 5 & 0 & -7 \end{pmatrix}$ usando el desarrollo por cofactores de dos filas diferentes y por una de sus columnas.

Usando la fila 2, tenemos

$$\begin{aligned}
\det A &= 0A_{21} + 4A_{22} - 2A_{23} + 0A_{24} \\
&= 4(-1)^{2+2}\begin{vmatrix} 2 & 0 & 3 \\ -3 & 1 & 2 \\ \mathbf{0} & \mathbf{0} & \mathbf{-7} \end{vmatrix} - 2(-1)^{2+3}\begin{vmatrix} 2 & -1 & 3 \\ -3 & 0 & 2 \\ \mathbf{0} & \mathbf{5} & \mathbf{-7} \end{vmatrix} \\
&= 4(\mathbf{-7})(-1)^{3+3}\begin{vmatrix} 2 & 0 \\ -3 & 1 \end{vmatrix} + 2\left\{\mathbf{5}(-1)^{3+2}\begin{vmatrix} 2 & 3 \\ -3 & 2 \end{vmatrix} \mathbf{-7}(-1)^{3+3}\begin{vmatrix} 2 & -1 \\ -3 & 0 \end{vmatrix}\right\} \\
&= -28(2+0) + 2\{-5(4+9) - 7(0-3)\} \\
&= -56 + 2\{-65 + 21\} = -144
\end{aligned}$$

[7]El cálculo del determinante de una matriz $n\times n$ implica el cálculo del determinante de $n\cdot(n-1)\cdot(n-2)\cdot\ \cdots\ \cdot 3 = \frac{n!}{2}$ matrices 2×2, lo cual requiere al menos $2\cdot\frac{n!}{2} = n!$ multiplicaciones (además de un número similar de sumas).

Usando la fila 4, tenemos

$$\begin{aligned}
\det A &= 0A_{41}+5A_{42}+0A_{43}-7A_{44}\\
&= 5(-1)^{4+2}\begin{vmatrix} 2 & 0 & 3\\ \mathbf{0} & \mathbf{-2} & \mathbf{0}\\ -3 & 1 & 2\end{vmatrix} - 7(-1)^{4+4}\begin{vmatrix} \mathbf{2} & -1 & 0\\ \mathbf{0} & 4 & -2\\ \mathbf{-3} & 0 & 1\end{vmatrix}\\
&= 5(\mathbf{-2})(-1)^{2+2}\begin{vmatrix} 2 & 3\\ -3 & 2\end{vmatrix} - 7\left\{\mathbf{2}(-1)^{1+1}\begin{vmatrix} 4 & -2\\ 0 & 1\end{vmatrix}\mathbf{-3}(-1)^{3+1}\begin{vmatrix} -1 & 0\\ 4 & -2\end{vmatrix}\right\}\\
&= -10(4+9)-7\{2(4+0)-3(2-0)\}\\
&= -130-7\{8-6\}=-144
\end{aligned}$$

Finalmente, usando la columna 1, tenemos

$$\begin{aligned}
\det A &= 2A_{11}+0A_{21}-3A_{31}+0A_{41}\\
&= 2(-1)^{1+1}\begin{vmatrix} 4 & -2 & 0\\ \mathbf{0} & \mathbf{1} & \mathbf{2}\\ 5 & 0 & -7\end{vmatrix} - 3(-1)^{3+1}\begin{vmatrix} -1 & \mathbf{0} & 3\\ 4 & \mathbf{-2} & 0\\ 5 & \mathbf{0} & -7\end{vmatrix}\\
&= 2\left\{\mathbf{1}(-1)^{2+2}\begin{vmatrix} 4 & 0\\ 5 & -7\end{vmatrix} + \mathbf{2}(-1)^{2+3}\begin{vmatrix} 4 & -2\\ 5 & 0\end{vmatrix}\right\} - 3(\mathbf{-2})(-1)^{2+2}\begin{vmatrix} -1 & 3\\ 5 & -7\end{vmatrix}\\
&= 2\{-28-0-2(0+10)\}+6(7-15)\\
&= 2(-48)+6(-8)=-144
\end{aligned}$$

□

El Teorema 12, al igual que el ejemplo anterior, nos sugiere que debemos calcular los determinantes utilizando las filas o columnas con mayor número de ceros y así, reducir el número de cofactores a calcular. Por ejemplo, observemos que si la matriz es diagonal o triangular, como la matriz B del Ejemplo 29, el cálculo del determinante se simplifica enormemente.

Teorema 13 [*Determinante de una matriz triangular o diagonal*].
El determinante de una matriz diagonal o triangular es el producto de los elementos de su diagonal. En otras palabras, si $A_{(n)} = (a_{ij})$ es una matriz diagonal o triangular de tamaño $n \times n$, entonces

$$\det A_{(n)} = a_{11}a_{22}\dots a_{nn}. \tag{3.3}$$

Demostración: Usando la inducción matemática sobre el tamaño de la matriz, hagamos la demostración para una matriz triangular inferior. Si $A_{(2)}$ es una matriz triangular inferior 2×2, usando el resultado del Ejemplo 27, tenemos

$$\det A_{(2)} = \begin{vmatrix} a & 0\\ c & d\end{vmatrix} = ad-0c=ad.$$

Supongamos que (3.3) es válido para matrices triangulares inferiores de tamaño $(n-1)\times(n-1)$, es decir,

$$\det A_{(n-1)} = \begin{vmatrix} a_{11} & 0 & \cdots & 0\\ a_{21} & a_{22} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{n-1\ n-1}\end{vmatrix} = a_{11}a_{22}\cdot\ldots\cdot a_{n-1n-1}.$$

Demostremos que (3.3) es válido para matrices triangulares inferiores de tamaño $n\times n$. Usando la hipótesis de inducción para calcular el determinante de A por su desarrollo por cofactores de la primera fila, tenemos

$$\det A_{(n)} = \begin{vmatrix} a_{11} & 0 & \cdots & 0\\ a_{21} & a_{22} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & 0 & \cdots & 0\\ a_{31} & a_{33} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} = a_{11}(a_{22}\dots a_{nn}).$$

De manera similar, probamos que (3.3) es válido para matrices triangulares superiores. El resultado (3.3) para matrices diagonales lo obtenemos recordando que una matriz diagonal es una matriz triangular. $\square$

Este resultado es importante tanto para nuestro propósito de desarrollar una forma eficiente de calcular un determinante, como para la exploración y demostración de sus propiedades más usadas, por dos razones. Primero, como vimos en la Sección 7, toda matriz se puede reducir a una forma escalonada equivalente mediante operaciones elementales entre filas; y segundo, las propiedades básicas de los determinantes, las cuales presentamos en los siguientes teoremas, permiten establecer una relación bastante simple entre el determinante de una matriz y el determinante de su forma escalonada equivalente, la cual también es una matriz triangular.

Teorema 14 [*Propiedades básicas de los determinantes*].
Dada una matriz A de tamaño $n \times n$

1. Si A tiene una fila (o columna) de ceros, el determinante de A es 0.

2. Si la matriz B se obtiene al intercambiar dos filas (o columnas) de A, entonces $\det B = -\det A$.

3. Si la matriz A tiene dos filas (o columnas) iguales, entonces su determinante es 0.

4. Si la matriz B se obtiene de A al multiplicar una fila (o columna) por un escalar λ, entonces $\det B = \lambda \det A$.

5. Si A, B y C son matrices iguales excepto en la i-ésima fila de tal forma que la i-ésima fila (o columna) de C es la suma de las correspondientes i-ésimas filas (o columnas) de A y B, entonces $\det C = \det A + \det B$.

6. Si la matriz B se obtiene de A al sumar un múltiplo de una fila (o columna) a otra fila (o columna), entonces $\det B = \det A$.

Demostración: Realizaremos las pruebas de estas propiedades en términos de filas. Las pruebas directas de ellas, en términos de columnas, son completamente análogas o son resultados inmediatos al utilizar el Corolario 12.1.

1. Si la i-ésima fila de A es de sólo ceros, al desarrollar el determinante por cofactores utilizando esta fila, obtenemos
$$\det A = 0A_{i1} + 0A_{i2} + \ldots + 0A_{in} = 0$$

2. Haremos la prueba por inducción sobre el tamaño de la matriz. Podemos ver que, en matrices 2×2, la propiedad es válida.
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc = -(cb - da) = -\begin{vmatrix} c & d \\ a & b \end{vmatrix}.$$
Supongamos que el resultado es válido para matrices de tamaño $(n-1) \times (n-1)$. Si la matriz B se obtiene de A al intercambiar las filas i y j, para $r \neq i, j$, la fila r de A es igual a la fila r de B, los menores correspondientes a la r-ésima fila de las matrices A y B son de tamaño $(n-1) \times (n-1)$ y los menores de B se obtienen intercambiando dos filas en los correspondientes menores de A; por lo tanto, los cofactores de B tienen signo contrario a los correspondientes cofactores de A por la hipótesis de inducción. Así, que
$$\begin{aligned} \det A &= a_{r1}A_{r1} + a_{r2}A_{r2} + \ldots + a_{rn}A_{rn} \\ &= b_{r1}(-B_{r1}) + b_{r2}(-B_{r2}) + \ldots + b_{rn}(-B_{rn}) \\ &= -[b_{r1}B_{r1} + b_{r2}B_{r2} + \ldots + b_{rn}B_{rn}] \\ &= -\det B, \end{aligned}$$

3. Supongamos que las filas i y j son iguales. Al intercambiar estas dos filas, por la propiedad anterior, el determinante cambia de signo. Pero al intercambiar estas dos filas, obtenemos la misma matriz , así que el determinante es el mismo. Como el único número que es igual a su opuesto es el 0, el determinante de la matriz debe ser 0.

4. Si B es la matriz que se obtiene de multiplicar la i-ésima fila de A por λ, al calcular el determinante de B utilizando esta fila, tenemos que

$$\begin{aligned}\det B &= b_{i1}B_{i1} + b_{i2}B_{i2} + \cdots + b_{in}B_{in} \\ &= \lambda a_{i1}A_{i1} + \lambda a_{i2}A_{i2} + \cdots + \lambda a_{in}A_{in} \\ &= \lambda(a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}) \\ &= \lambda \det A.\end{aligned}$$

5. Al calcular el determinante de C utilizando la i-ésima fila, como $c_{ij} = a_{ij} + b_{ij}$, para $j = 1, 2, \ldots, n$, tenemos

$$\begin{aligned}\det C &= c_{i1}C_{i1} + c_{i2}C_{i2} + \cdots + c_{in}C_{in} \\ &= (a_{i1} + b_{i1})C_{i1} + (a_{i2} + b_{i2})C_{i2} + \cdots + (a_{in} + b_{in})C_{in} \\ &= (a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}) + (b_{i1}C_{i1} + b_{i2}C_{i2} + \cdots + b_{in}C_{in}) \\ &= (a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}) + (b_{i1}B_{i1} + b_{i2}B_{i2} + \cdots + b_{in}B_{in}) \\ &= \det A + \det B\end{aligned}$$

   ya que los cofactores de A, B y C correspondientes a la i-ésima fila son iguales.

6. Si la matriz B se obtiene de A al sumarle λ veces la j-ésima fila a la i-ésima fila, entonces

$$\begin{aligned}\det B &= b_{i1}B_{i1} + b_{i2}B_{i2} + \cdots + b_{in}B_{in} \\ &= (a_{i1} + \lambda a_{j1})B_{i1} + (a_{i2} + \lambda a_{j2})B_{i2} + \cdots + (a_{in} + \lambda a_{jn})B_{in} \\ &= (a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}) + \lambda(a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in}) \\ &= \det A + \lambda 0 \\ &= \det A\end{aligned}$$

   ya que los cofactores de las i-ésimas filas de A y B son los mismos y que

$$(a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in})$$

   corresponde al cálculo del determinante de una matriz que tiene las filas i y j iguales, utilizando la i-ésima fila. □

Con los resultados del teorema anterior, podemos conocer el determinante de las matrices elementales.

Corolario 14.1

Sea E una matriz elemental $n \times n$.

1. Si E es de Tipo Permutación, entonces $\det E = -1$.
2. Si E es de Tipo Escalamiento (resulta de multiplicar una fila de I por c), entonces $\det E = c$.
3. Si E es de Tipo Eliminación, entonces $\det E = 1$. □

Teniendo en cuenta que pre-multiplicar una matriz por una matriz elemental es equivalente a aplicarle a la matriz la operación elemental correspondiente, tenemos el siguiente resultado.

Corolario 14.2
Sean E y A matrices de igual tamaño, donde E es una matriz elemental. Entonces

$$\det(EA) = (\det E)(\det A).$$

□

Como dijimos anteriormente, las propiedades de los determinantes presentadas hasta ahora y el Algoritmo de Eliminación de Gauss para escalonar una matriz nos permiten proponer una forma eficiente de calcular un determinante. Recordemos que, dada una matriz de tamaño $n \times n$, podemos encontrar su forma escalonada mediante operaciones elementales entre filas de Tipo Eliminación y Permutación solamente, lo cual es equivalente a decir que existen $E_1,\ E_2\ \ldots,\ E_k$, matrices elementales de Tipo Eliminación y Permutación, tales que

$$U = E_k \ldots\ E_2 E_1 A,$$

donde U es una matriz escalonada (triangular superior) de tamaño $n \times n$. Utilizando este resultado, el siguiente teorema propone la forma eficiente de calcular un determinante antes anunciada.

Teorema 15 *[Cálculo del determinante de una matriz].*
Dada A, una matriz de tamaño $n \times n$, sea $U = (u_{ij})$ la matriz triangular superior obtenida mediante la aplicación de operaciones elementales Tipo Eliminación y Permutación a la matriz A. Entonces,

$$\det A = (-1)^p u_{11} u_{22} \cdots u_{nn},$$

donde p es el número de operaciones elementales Tipo Permutación utilizadas para obtener la matriz U a partir de la matriz A.

Demostración: Sean $E_1,\ E_2, \ldots,\ E_k$ las matrices elementales correspondientes a las operaciones elementales aplicada a A para obtener U. Entonces,

$$U = E_k \cdots\ E_2 E_1 A.$$

Por el Corolario 14.2,

$$\det U = (\det E_k) \cdots (\det E_2)(\det E_1)(\det A).$$

Como $\det E_i = -1$ para las p matices elementales Tipo Permutación y $\det E_i = 1$ para el resto de matrices elementales (Corolario 14.1), tenemos

$$\det U = (-1)^p \det A.$$

Finalmente, como U es una matriz triangular superior, por el Teorema 13, tenemos

$$\det A = (-1)^p \det U = (-1)^p u_{11} u_{22} \cdots u_{nn}.$$

□

Ejemplo 31. Calculemos el determinante de $A = \begin{pmatrix} 0 & 6 & -1 & -5 \\ -3 & 2 & 3 & 0 \\ -6 & -5 & 3 & 8 \\ -6 & 4 & 0 & 2 \end{pmatrix}$ y de $B = \begin{pmatrix} 8 & 0 & 16 \\ -6 & 0 & 9 \\ 4 & -1 & -6 \end{pmatrix}$.

Aplicando sucesivamente las operaciones elementales $F_1 \longleftrightarrow F_4$, $F_2 - \frac{1}{2}F_1 \longrightarrow F_2$, $F_3 - F_1 \longrightarrow F_3$, $F_2 \longleftrightarrow F_3$, $F_4 + \frac{2}{3}F_2 \longrightarrow F_4$ y $F_4 - \frac{1}{3}F_3 \longrightarrow F_4$ a la matriz A, obtenemos una forma escalonada equivalente

$$U = \begin{pmatrix} -6 & 4 & 0 & 2 \\ 0 & -9 & 3 & 6 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & -2/3 \end{pmatrix}.$$

Por lo tanto, $\det A = (-1)^2 \det U = (-6)(-9)3(-2/3) = -108$, puesto que hubo dos operaciones elementales Tipo Permutación.

De la misma forma, aplicando sucesivamente las operaciones elementales $F_2 + \frac{3}{4}F_1 \longrightarrow F_2$, $F_3 + \frac{1}{2}F_1 \longrightarrow F_3$, $F_2 \longleftrightarrow F_3$ a la matriz B, obtenemos una forma escalonada equivalente

$$U = \begin{pmatrix} 8 & 0 & -16 \\ 0 & -1 & 2 \\ 0 & 0 & -3 \end{pmatrix}.$$

Por lo tanto, $\det B = (-1)^1 \det U = (-1)8(-1)(-3) = -24$, ya que hubo una operación elemental Tipo Permutación. □

El resultado anterior nos permite probar una relación teórica importante entre determinantes e invertibilidad de matrices, la cual consignamos en el siguiente corolario.

Corolario 15.1
Sea A una matriz de tamaño $n \times n$. La matriz A es invertible, si y sólo si, $\det A \neq 0$.

Demostración: Sea U la matriz triangular superior obtenida mediante la aplicación de operaciones elementales Tipo Eliminación y Permutación a la matriz A. Por el Teorema 15, $\det A \neq 0$, si y solo si, $u_{ii} \neq 0$ para todo i, lo cual ocurre, si y solo si, U tiene n pivotes, lo cual, por el Teorema 7, es equivalente a que U es invertible. □

Como podemos verificar directamente (ver Ejemplo 32), el determinante no tiene la propiedad distributiva con la suma de matrices, pero si con el producto de matrices. Además, tiene una propiedad particular respecto de la multiplicación por escalar de matrices, como lo expresamos en el siguiente teorema.

Teorema 16
Si A y B son matrices $n \times n$ y α un número real (escalar), entonces

1. $\det(\alpha A) = \alpha^n \det A$.
2. $\det(AB) = \det A \det B$.
3. $\det(A^m) = (\det A)^m$.

Demostración:

1. Al multiplicar una fila por un escalar α, por el Teorema 14 el determinante de la nueva matriz es α veces el determinante de la matriz original y multiplicar una matriz por un escalar α es equivalente a multiplicar cada una de sus filas por α, así que el determinante de αA es el determinante de A multiplicado por α, n veces; es decir, $\det(\alpha A) = \alpha^n \det A$.

2. Separemos la demostración en dos casos, según la invertibilidad de la matriz A. Sea A una matriz no invertible. Entonces, por el Corolario 7.1, AB tampoco es invertible, por el Corolario 15.1, $\det A = 0$ y $\det(AB) = 0$; asi que $\det(AB) = \det A \det B$.
Sea A una matriz invertible. Por el Teorema 10, existen matrices elementales $E_1, E_2, \ldots, E_k$ tales que
$$E_k \cdots E_2 E_1 = A,$$
por tanto,
$$\det(AB) = \det((E_k \cdots E_2 E_1)B) = \det(E_k(\cdots(E_2(E_1 B))\cdots)).$$
Aplicando el Corolario 14.2 repetidamente, obtenemos que
$$\det(AB) = (\det E_k)\cdots(\det E_2)(\det E_1)(\det B)$$
y como, por el mismo Corolario 14.2,
$$\det A = (\det E_k)\cdots(\det E_2)(\det E_1),$$
entonces, $\det(AB) = (\det A)(\det B)$.

3. Usando la inducción matemática sobre la potencia m, por el resultado anterior,

$$\det(A^2) = \det(AA) = (\det A)(\det A) = (\det A)^2.$$

Supongamos que el resultado es válido para $m-1$; es decir, $\det(A^{m-1}) = (\det A)^{m-1}$. Entonces, por el resultado anterior y la hipótesis de inducción, tenemos

$$\det(A^m) = \det(AA^{m-1}) = (\det A)\det(A^{m-1}) = (\det A)(\det A)^{m-1} = (\det A)^m.$$

□

Corolario 16.1
Si A es una matriz invertible, entonces

$$\det(A^{-1}) = \frac{1}{\det A}.$$

Demostración: Como $AA^{-1} = I$, por el Resultado 2 del teorema anterior, $(\det A)\det(A^{-1}) = 1$, de donde obtenemos la igualdad deseada despejando $\det(A^{-1})$. □

Es importante que notemos que, aunque $AB \neq BA$, $\det(AB) = \det(BA)$ y que $\det(A+B) \neq \det A + \det B$ en general.

Ejemplo 32. Dadas las matrices $A = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 2 & 1 \\ 3 & 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & -2 \\ 0 & 3 & 1 \end{pmatrix}$, calcule $\det A$, $\det B$, $\det(2A)$, $\det(AB)$, $\det(A+B)$.

Por el tamaño de la matrices y la cantidad de ceros en una misma fila o columna, calculemos el determinante de A y B usando su desarrollo por cofactores

$$|A| = 3(-1)^{3+1}\begin{vmatrix} 0 & 1 \\ 2 & 1 \end{vmatrix} = 3(0-2) = -6$$

$$|B| = 3(-1)^{3+2}\begin{vmatrix} 2 & 0 \\ 1 & -2 \end{vmatrix} = -3(-4-0) = 12$$

Apliquemos las propiedades de los determinantes presentadas para calcular los determinantes de $2A$ y AB

$$|2A| = 2^3|A| = 8(-6) = -48 \quad \text{y} \quad |AB| = |A||B| = (-6)12 = -72$$

Como el determinante no es distributivo con la suma de matrices, debemos calcular $A+B$ para calcular su determinante.

$$|A+B| = \begin{vmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ 3 & 3 & 1 \end{vmatrix} = 1(-1)^{1+1}\begin{vmatrix} 2 & -1 \\ 3 & 1 \end{vmatrix} + 1(-1)^{1+3}\begin{vmatrix} 1 & 2 \\ 3 & 3 \end{vmatrix} = (2+3) + (3-6) = 2$$

Observemos que $|A+B| \neq |A| + |B|$. □

Otra propiedad teórica sobre los determinantes está relacionada con la llamada matriz adjunta, la cual permite escribir cada una de las componentes de la matriz inversa de una matriz invertible en término de determinantes. Sobra advertir que esta propiedad, en general, no es un elemento de cálculo, por el alto costo computacional involucrado en el cálculo de los determinantes, asi sea usando la eliminación de Gauss.

Definición 12 [*Adjunta*]. Dada A una matriz de tamaño $n \times n$, definimos la *matriz de cofactores* de A como la matriz cuya componente (i, j) es el cofactor A_{ij} y definimos $Adj(A)$, la *matriz adjunta de A*, como la

transpuesta de la matriz de cofactores. En otras palabras,

$$Adj(A) = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

Ejemplo 33. Dada la matriz $A = \begin{pmatrix} 2 & 3 & 0 \\ 1 & -1 & -2 \\ 7 & 1/2 & 4 \end{pmatrix}$, calculemos la matriz adjunta de A.

Para calcular la adjunta de A, tenemos que calcular los 9 cofactores de A:

$$\begin{array}{lll} A_{11} = -3 & A_{12} = -18 & A_{13} = 15/2 \\ A_{21} = -12 & A_{22} = 8 & A_{23} = 20 \\ A_{31} = 6 & A_{32} = 4 & A_{33} = -5 \end{array}$$

Así que

$$Adj(A) = \begin{pmatrix} -3 & -12 & 6 \\ -18 & 8 & 4 \\ 15/2 & 20 & -5 \end{pmatrix}.$$

□

Es fácil ver que al multiplicar la matriz $Adj(A)$ por la matriz A, obtenemos $(\det A)I_3$. Este resultado es válido en general, como lo expresa el siguiente teorema.

Teorema 17
Si A es una matriz cuadrada, entonces

$$A \; Adj(A) = (\det A)I = Adj(A) \; A.$$

Demostración: Observemos que, en el producto

$$A \; Adj(A) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

la componente (i, j) es

$$fil_i(A) \cdot col_j(AdjA) = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} A_{j1} \\ A_{j2} \\ \vdots \\ A_{jn} \end{pmatrix} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}.$$

Este producto es el desarrollo por cofactores del determinante de la matriz A', que se obtiene de la matriz A al reemplazar la fila j por la fila i. Cuando $i \neq j$, la matriz A' tiene dos filas iguales, por tanto su determinante es 0. Así que la componente (i, j) de $A \; Adj(A)$ es 0, cuando $i \neq j$. Pero, cuando $i = j$, este producto es el desarrollo por cofactores del determinante de la matriz $A' = A$ usando la fila i. La demostración de la segunda igualdad del teorema es similar a la anterior. □

Este resultado nos permite una caracterización de la inversa de una matriz cuando $\det A \neq 0$, lo cual resumimos en el siguiente corolario.

Corolario 17.1
Si A es una matriz invertible, entonces

$$A^{-1} = \frac{1}{\det A} Adj(A)$$

Demostración: Si A es una matriz invertible, $\det A \neq 0$; por lo tanto, usando las propiedades del producto de matrices y el teorema anterior, tenemos

$$A\left(\frac{1}{\det A} Adj(A)\right) = \frac{1}{\det A}\left(A\ Adj(A)\right) = \frac{1}{\det A}(\det A)I = I.$$

□

Ejemplo 34. Calculemos la componente $(3,2)$ de la inversa de la matriz $A = \begin{pmatrix} -2 & 0 & 1 & 0 \\ 0 & -1 & 0 & -2 \\ 4 & 0 & -7 & -1 \\ 0 & 3 & 0 & 1 \end{pmatrix}$.

Sea $A^{-1} = (\alpha_{ij})$. Por el teorema anterior y la definición de matriz adjunta, $\alpha_{32} = \dfrac{A_{23}}{|A|}$. Calculemos el determinante de A usando operaciones elementales entre filas para calcular una forma escalonada equivalente a A. En este caso, aplicando sucesivamente las operaciones elementales $F_3 + 2F_1 \longrightarrow F_3$ y $F_4 + 3F_2 \longrightarrow F_4$, obtenemos la matriz escalonada

$$U = \begin{pmatrix} -2 & 0 & 1 & 0 \\ 0 & -1 & 0 & -2 \\ 0 & 0 & -5 & -1 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Por lo tanto, como no hubo permutación de filas, $|A| = |U| = -2(-1)(-5)5 = -50$.

Ahora, para calcular A_{23}, el cofactor $(2,3)$ de A, debemos calcular el determinante de $M_{23} = \begin{pmatrix} -2 & 0 & 0 \\ 4 & 0 & -1 \\ 0 & 3 & 1 \end{pmatrix}$, el menor $(3,2)$ de A.

$$A_{23} = (-1)^{2+3}|M_{23}| = -(-2)\begin{vmatrix} 0 & -1 \\ 3 & 1 \end{vmatrix} = 2[0(1) - 3(-1)] = 6.$$

En consecuencia, la componente $(3,2)$ de la inversa de A es $\alpha_{32} = \dfrac{A_{23}}{|A|} = \dfrac{6}{-50} = -0{,}12$. □

3.10. Ejercicios

1. Dadas las siguientes matrices,

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & -3 \\ -5 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1/3 & 0 & 0 & 0 \\ 5 & \sqrt{2} & 0 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1/3 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$D = \begin{pmatrix} \sqrt{2} & -1 & 1 & 0 \\ 0 & \sqrt{2} & 7 & 2 \\ 0 & 0 & \sqrt{2} & -5 \end{pmatrix}, \qquad E = \begin{pmatrix} 4 & 0 \\ 5 & -2 \\ -2 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad F = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix},$$

clasifíquelas en triangulares inferiores, triangulares superiores, escalares o diagonales.

2. Que información de cada uno de los siguientes tipos de matrices es suficiente y necesario guardar para reconstruir toda la matriz?

   a) Triangular inferior b) Diagonal c) Escalar d) Simétrica.

3. Demuestre el Teorema 2.

4. Dadas las matrices A, B, C y D, identifique las expresiones matriciales que están bien definidas y calcúlelas.

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 5 & 0 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 1/3 & 0 & 2 \\ 2 & -3 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 5 \\ 0 & -2 \\ -1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} -1 & 2 & -2 \\ 4 & -3 & -1 \end{pmatrix}$$

   a) $2(A - B) + 3A + 2B$
   b) $-3(A + B - C) - 2D$
   c) $-2(ACB + DCB)$
   d) $3[-5(A + 2B - D)] - 2[1/2(A + 2B - D)]$
   e) $A^T B + A^T D + C^T D$
   f) $(A^T D)^{-1} - C^{-1}$

5. Analice el Ejemplo 8 del texto.

6. Para las matrices A y D del Ejercicio 4, resuelva la siguiente ecuación matricial.

$$A - 2X + 3D = -A - 5D.$$

7. Complete la demostración del Teorema 3.

8. Para las matrices B y C del Ejercicio 4, calcule

   a) La componente (1,2) de BC. b) La segunda fila de BC. c) La tercera columna de BC.

9. Cuáles de las siguientes matrices son invertibles y diga por qué.

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 5 & 0 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 1/3 & 0 & 1 \\ 0 & -3 & -6 \\ 0 & 1 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1/3 & 0 & 2 & -1 \\ 0 & -3 & -1 & 2 \\ 0 & 0 & 2 & -5 \\ 0 & 0 & -4 & 3 \end{pmatrix}.$$

10. Complete la demostración del Teorema 6.

11. Dadas las matrices invertibles A y B y el vector $\mathbf{b}$

$$A = \begin{pmatrix} -1 & \sqrt{2} & 1 \\ 5 & 1 & -3 \\ 0 & 3 & -7 \end{pmatrix}, \quad B = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & -1 & -1/2 \\ 0 & 2 & 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ -5 \end{pmatrix},$$

   calcule

$$[AB + 3A]^T \left(\frac{1}{2} B^T A^T \right)^{-1} \mathbf{b}$$

12. Demuestre que si A es invertible y $AB = AC$, entonces $B = C$.

13. Demuestre que AA^T es una matriz simétrica para cualquier matriz A.

14. Complete la demostración del Teorema 8.

15. Qué tipo de matriz es la suma de dos matrices diagonales? El producto?

16. Qué tipo de matriz es el producto por un escalar de una matriz diagonal? Su transpuesta? Su inversa?.

17. Qué tipo de matriz es el producto por un escalar de una matriz triangular? Su transpuesta? Su inversa?

18. Cómo deben ser los elementos de una matriz diagonal para ser invertible? En este caso, cual es la matriz inversa?

19. Para qué valor o valores de α, A^{-1} existe?

$$a)A = \begin{pmatrix} 1 & -2 & 1 \\ 5 & \alpha & -3 \end{pmatrix}, \qquad b)A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & -3 & -6 \\ \alpha & 1 & 2 \end{pmatrix}, \qquad A = \begin{pmatrix} 2 & 0 & 2 & -1 \\ 0 & -3 & -1 & 2 \\ -1 & 0 & 2 & -5 \\ 0 & 3 & -4 & \alpha \end{pmatrix}.$$

20. Demuestre que si la matriz A es invertible y $AB = 0$, entonces $B = 0$.

21. Decida si cada una de las siguientes proposiciones es falsa o verdadera y diga por qué.

    a) Si las columnas de una matriz forman un conjunto de vectores *l.i.*, la matriz es invertible.
    b) La suma de dos matrices invertibles es una matriz invertible.
    c) Si la solución de un sistema de ecuaciones lineales homogéneo es única, la matriz de coeficientes del sistema es invertible.
    d) Cualquier múltiplo diferente de cero de una matriz invertible es invertible.
    e) Si la suma de dos matrices es invertible, cada una de las matrices es invertible.
    f) Si el producto de dos matrices es invertible, cada una de las matrices es invertible.
    g) El producto de dos matrices cuadradas siempre es invertible.
    h) Si una matriz es invertible, cualquier forma escalonada equivalente también es invertible.
    i) Si la forma escalonada equivalente de una matriz es invertible, la matriz también es invertible.
    j) Si el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución única, la matriz A es invertible.
    k) Si la matriz A es invertible, el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución única.
    l) Si la suma de las columnas de una matriz es igual al vector 0, la matriz no es invertible.
    m) Si el sistema $A\mathbf{x} = \mathbf{b}$ tiene infinitas soluciones, la matriz A no es invertible.
    n) Si el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución única y A es cuadrada, la matriz A es invertible.
    ñ) Para concluir que una matriz es invertible, es necesario calcular su inversa.
    o) Si la matriz aumentada de un sistema de ecuaciones lineales es invertible, el sistema es inconsistente.

22. Decimos que una matriz A es *idempotente*, si y solo si, $A^2 = A$ y decimos que una matriz A es *nilpotente*, si y solo si, alguna potencia de la matriz es cero; es decir, $A^k = O$ para algún valor de $k \in N$. Decida si cada una de las siguientes proposiciones es falsa o verdadera y diga por qué.

    a) Toda matriz idempotente es invertible.
    b) Ninguna matriz nilpotente es invertible.
    c) La matriz idéntica es una matriz idempotente.
    d) Una matriz triangular con los elementos de la diagonal iguales a cero es una matriz nilpotente.
    e) Si $A^4 = O$, la matriz A es nilpotente.
    f) Si $A^4 = O$, la matriz A^2 es nilpotente.
    g) Si $A^4 = A^2$, la matriz A es idempotente.
    h) Si $A^4 = A^2$, la matriz A^2 es idempotente.
    i) Si A es idempotente, A^2 también es idempotente.

23. Decimos que una matriz A cuadrada es *ortogonal*, si y solo si, las columnas de la matriz son vectores unitarios y ortogonales entre si. Decida si cada una de las siguientes proposiciones es falsa o verdadera y diga por qué.

    a) Si A es una matriz ortogonal, entonces $A^T A = I$.

    b) Toda matriz ortogonal es invertible.

    c) Si A es una matriz ortogonal, A^T también es ortogonal.

    d) Una matriz ortogonal simétrica es una matriz idempotente.

    e) La matriz idéntica es ortogonal.

    f) La inversa de una matriz ortogonal es su transpuesta.

    g) Si A es una matriz ortogonal, la solución del sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$ es $\mathbf{x} = A^T\mathbf{b}$.

24. Dadas las siguientes matrices, identifique el tipo de matriz elemental y escriba la correspondiente operación elemental.

$$E_1 = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 1 & 0 & -5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

25. Para cada una de las operaciones elementales dadas, escriba la correspondiente matriz elemental y su matriz inversa.

    a) $F_2 + \frac{3}{2}F_1 \to F_2$ b) $-5F_3 \to F_3$, c) $F_1 \leftrightarrow F_3$,

26. Al escalonar la matriz A, se aplicaron las operaciones elementales $F_2 + \frac{3}{2}F_1 \to F_2$, $F_3 + F_2 \to F_3$, y se obtuvo la matriz $\begin{pmatrix} 2 & -4 & 0 \\ 0 & -7 & 7 \\ 0 & 0 & 12 \end{pmatrix}$.

    a) Calcule la factorización LU de la matriz A.

    b) Resuelva $A\mathbf{x} = \mathbf{b}$ para $\mathbf{b}^T = (-1 \;\; 0 \;\; 6)$.

    c) Resuelva $A\mathbf{z} = \mathbf{d}$ para $\mathbf{d} = 3\mathbf{x}$.

    d) Resuelva $A^T\mathbf{y} = \mathbf{c}$ para $\mathbf{c}^T = (2 \;\; 0 \;\; -3)$.

    e) Existe $\det A$?

27. Si A y B son matrices 5×5 tales que $\det A = -3$ y $\det B = 1/2$, calcule, de ser posible,

    a) $\det(A^{-1})$. b) $\det(2B)$. c) $\det(AB)$. d) $\det(A + B)$.
    e) $\det(A^2)$. f) $\det(A^T)$. g) $\det[(3A)^{-1}]$. h) $\det[Adj(A)]$.
    i) $\det[(2B)^{-1}A(Adj(A))B^2]$

28. Si $A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ y $\det A = -\frac{1}{5}$, calcule, de ser posible, el determinante de las siguientes matrices

a) $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ b) $\begin{pmatrix} 5a_1 & a_2 & a_3 \\ 5b_1 & b_2 & b_3 \\ 5c_1 & c_2 & c_3 \end{pmatrix}$ c) $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1-2a_1 & c_2-2a_2 & c_3-2a_3 \end{pmatrix}$

d) $\begin{pmatrix} a_1 & b_1 & c_1 \\ -2a_2 & -2b_2 & -2c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ e) $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ 0 & 0 & 0 \end{pmatrix}$ f) $\begin{pmatrix} a_1 & 3a_1+b_1 & c_1 \\ a_2 & a_2+b_2 & c_2 \\ a_3 & a_3+3b_3 & c_3 \end{pmatrix}$

g) $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ 3c_1-2a_1 & 3c_2-2a_2 & 3c_3-2a_3 \end{pmatrix}$ h) $\begin{pmatrix} a_1 & a_3 & a_2 \\ b_1 & b_3 & b_2 \\ c_1 & c_3 & c_2 \end{pmatrix}$ i) $\begin{pmatrix} a_1 & 5a_1-a_2 & a_3 \\ b_1 & 5b_1-b_2 & b_3 \\ c_1 & 5c_1-c_2 & c_3 \end{pmatrix}$

29. Al escalonar la matriz A, se aplicaron las operaciones elementales $F_2+2F_1 \to F_2$, $F_3-\frac{5}{2}F_1 \to F_3$ y $F_3+\frac{10}{11}F_2 \to F_3$ y se obtuvo la matriz $\begin{pmatrix} -2 & 4 & 4 \\ 0 & 11 & 6 \\ 0 & 0 & -17/11 \end{pmatrix}$.

a) Calcule $\det A$. b) Existe A^{-1}?

c) Calcule la factorizacion LU de A. d) Calcule $\det Adj(A)$.

30. Conteste falso o verdadero y diga por qué

a) Si las columnas de una matriz cuadrada son $l.i.$, el determinante de la matriz es diferente de cero.

b) Si la solución de un sistema de ecuaciones lineales homogéneo cuadrado es única, el determinante de la matriz de coeficientes del sistema puede ser cero.

c) El determinante de $3A$ es 3 veces el determinante de A.

d) Si el determinante del producto de dos matrices es cero, una de las matrices es cero.

e) Si el determinante del producto de dos matrices es cero, el determinante de una de las matrices es cero.

f) La factotización LU de una matriz requiere mas operaciones que el uso de la eliminación de Gauss para escalonar la matriz.

g) Si una matriz es invertible, siempre es posible calcular su factorización LU.

h) Si la matriz A es invertible, la matriz A^4 también es invertible.

i) Si la matriz A es invertible, su matriz adjunta también es invertible.

j) Si la suma de las columnas de una matriz es igual al vector 0, su determinante es cero.

k) Si el determinante de una matriz es cero, una columna es combinación lineal de las otras.

l) Si el determinante de la matriz de coeficientes de un sistema de ecuaciones lineales es cero, el sistema es inconsistente.

m) Toda matriz es el producto de matrices elementales.

n) Toda matriz cuadrada es el producto de matrices elementales.

ñ) Toda matriz elemental es invertible.

o) Si dos matrices tienen dos filas iguales, sus determinantes también son iguales.

p) Al pre-multiplicar una matriz por una matriz elemental, el determinante de la matriz no cambia.

31. Dados los números reales $\alpha_0, \alpha_1, \ldots, \alpha_n$, la matriz de la forma

$$V_n = \begin{pmatrix} 1 & \alpha_0 & \alpha_0^2 & \cdots & \alpha_0^n \\ 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^n \end{pmatrix}$$

se conoce como la matriz de *Vandermonde*. Para el caso particular $n = 2$, demuestre que V_2 es invertible, si y solo si, $\alpha_0, \alpha_1, \alpha_2$ son distintos [AYUDA: Demuestre que $\det V_2 = (\alpha 1 - \alpha_0)(\alpha_2 - \alpha_0)(\alpha_2 - \alpha_1)$].

Capítulo 4

ESPACIOS VECTORIALES

4.1. Introducción

Después de haber estudiado los vectores de R^n y las matrices, podemos notar cierta similitud entre las propiedades de sus operaciones básicas (suma y multiplicación por escalar) lo cual no es casual. En realidad, los vectores de R^n y las matrices son dos ejemplos muy importantes de aquellos conjuntos en los cuales hay definidas dos operaciones como la suma y la multiplicación por escalar, las cuales tienen ciertas propiedades (como las probadas para los vectores de R^n y las matrices). En este capítulo, estudiaremos este tipo de conjuntos, a los cuales llamaremos Espacios Vectoriales, con el fin de analizar, en forma general, las consecuencias o ventajas de tener un conjunto con estas dos operaciones, sin tener que demostrarlas para cada caso particular. Centraremos nuestra atención en el estudio de los espacios vectoriales de dimensión finita donde los escalares son números reales *(espacio vectorial real finito)*, ya que estos nos permiten visualizar una amplia gama de conjuntos dotados con este tipo de operaciones como si fuera el conjunto de vectores de R^n, para un n en particular.

Concretamente, en este capítulo, generalizaremos los conceptos de vector, combinación lineal, independencia lineal, conjuntos generados y generadores, norma, paralelismo y ortogonalidad, los cuales serán, en apariencia, iguales a los dados en el Capítulo 2. Con base en lo anterior, introduciremos los conceptos de base, de dimensión y de componentes de un vector en una base, lo cual nos permitirá analizar y hacer cálculos sobre cualquier espacio vectorial real finito como si se tratara de R^n, lo que resalta la importancia de conocer muy bien las propiedades de los vectores en R^n.

4.2. Definición y Propiedades Básicas

Definición 1 [*Espacio vectorial*]. Sea V un conjunto no vacío en el cual se han definido dos operaciones, llamadas *suma* y *multiplicación por escalar* (dados los elementos $\mathbf{u}$ y $\mathbf{v}$ de V y un escalar λ de R, la suma de $\mathbf{u}$ y $\mathbf{v}$ la denotamos $\mathbf{u}+\mathbf{v}$ y la multiplicación por escalar de λ por $\mathbf{u}$ la denotamos $\lambda\mathbf{u}$). Si las siguientes propiedades o axiomas se satisfacen para todo $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ de V y para todo par de escalares α y β de R, entonces se dice que V es un *espacio vectorial real* y sus elementos son llamados *vectores*.

1. $\mathbf{u}+\mathbf{v} \in V$. *Propiedad clausurativa para la suma.*

2. $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$. *Propiedad conmutativa para la suma.*

3. $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$. *Propiedad asociativa para la suma.*

4. Existe un único elemento $\mathbf{0} \in V$, tal que $\mathbf{u}+\mathbf{0} = \mathbf{u}$, para todo $\mathbf{u} \in V$. *Propiedad modulativa para la suma.*

5. Para cada $\mathbf{u} \in V$, existe un único elemento $-\mathbf{u} \in V$, tal que $\mathbf{u}+(-\mathbf{u}) = \mathbf{0}$. *Existencia del opuesto para la suma.*

6. $\alpha\mathbf{u} \in V$. *Propiedad clausurativa para la multiplicación por escalar.*
7. $\alpha(\mathbf{u}+\mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$. *Propiedad distributiva respecto la suma de vectores.*
8. $(\alpha+\beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$. *Propiedad distributiva respecto la suma de escalares.*
9. $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

Observemos que un espacio vectorial V se define como un conjunto de elementos en el cual están definidas dos operaciones *suma* y *producto por escalar* que satisfacen las 10 propiedades anteriores y que no se especifica la naturaleza de los elementos de V ni de las operaciones. Con frecuencia, las operaciones tendrán la naturaleza de la suma y multiplicación que nos es familiar, pero en otros casos no. Como lo mencionamos en la introducción, ya tenemos dos ejemplos de espacio vectorial real, los vectores de R^n con la suma y el producto por escalar definidos en el Capítulo 2 y el conjunto de las matrices de un tamaño dado con la suma y la multiplicación por escalar definidas en el Capítulo 3.

En los otros ejemplos que presentaremos, debemos describir o caracterizar los elementos del conjunto, especificar las operaciones y verificar que se satisfagan los 10 axiomas de la Definición 1. De los axiomas, por razones que entenderemos más adelante, es conveniente primero verificar los Axiomas 1 y 6 (propiedades clausurativas), luego los Axiomas 4 y 5 (axiomas de existencia) y luego los demás que son propiedades relacionadas con las operaciones, más que con los elementos o vectores del espacio vectorial.

En los Capítulos 2 y 3, cuando definimos el producto por escalar en los vectores de R^n y en las matrices, nos referimos a un escalar como un número real, pero en realidad podríamos pensar en otros conjuntos de escalares, por ejemplo, los números racionales o los números complejos. En la mayoría de los casos, los escalares serán números reales (espacios vectoriales reales o simplemente espacios vectoriales), y, mientras no digamos lo contrario, asi lo asumiremos.

Ejemplo 1. Sea $\mathcal{P}_2$ el conjunto de todos los polinomios de grado menor o igual a 2. Es decir,

$$\mathcal{P}_2 = \{p(x) = a_0 + a_1x + a_2x^2 : a_i \in R, i = 0,1,2\}.$$

Si definimos la suma y el producto por escalar como se hace habitualmente; esto es, dados

$$p(x) = a_0 + a_1x + a_2x^2 \quad \text{y} \quad q(x) = b_0 + b_1x + b_2x^2,$$

dos polinomios de grado menor o igual a 2, y λ, un escalar, entonces

$$\begin{aligned} p(x) + q(x) &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 \\ \lambda p(x) &= (\lambda a_0) + (\lambda a_1)x + (\lambda a_2)x^2. \end{aligned}$$

Es fácil verificar que al sumar dos polinomios de grado menor o igual a dos, obtenemos un polinomio cuyo grado es menor o igual a dos (Axioma 1) y que al multiplicar un polinomio de grado $k \leq 2$ por un escalar, obtenemos de nuevo un polinomio de grado k, a menos que el escalar por el que multipliquemos sea 0, en cuyo caso, el polinomio es $\mathbf{0}$ (el polinomio cuyos coeficientes son 0) (Axioma 6).

La existencia del vector nulo (Axioma 4) es clara, el vector nulo es el polinomio $\mathbf{0}$, y el opuesto para la suma de un polinomio $p(x) = a_0 + a_1x + a_2x^2$ es el polinomio cuyos coeficientes son los opuestos de los de $p(x)$, $-p(x) = -a_0 - a_1x - a_2x^2$ (Axioma 5). Los demás axiomas se verifican fácilmente. Verifiquemos el Axioma 9 y dejemos los demás como ejercicio para el lector. Tenemos que

$$\begin{aligned} \alpha[\beta p(x)] &= \alpha[\beta(a_0 + a_1x + a_2x^2)] \\ &= \alpha[\beta a_0 + \beta a_1x + \beta a_2x^2] \\ &= (\alpha\beta)a_0 + (\alpha\beta)a_1x + (\alpha\beta)a_2x^2 \\ &= (\alpha\beta)(a_0 + a_1x + a_2x^2) \\ &= (\alpha\beta)p(x). \end{aligned}$$

Por tanto, $\mathcal{P}_2$ es un espacio vectorial. □

Después de este ejemplo, podemos pensar en muchos más ejemplos similares: los polinomios de grado menor o igual a n, $\mathcal{P}_n = \{p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n : a_i \in R, i = 0, 1, 2, \ldots, n\}$, con operaciones suma y multiplicación por escalar definidas similarmente, para cualquier $n = 0, 1, 2, \ldots$.

Ejemplo 2. Observemos que el conjunto de polinomios de grado igual a n, con n fijo y $n \geq 1$, con las suma y la multiplicación por escalar definidas como antes, **no** es un espacio vectorial. En efecto, dicho conjunto no tiene el elemento neutro de la suma: el polinomio **0**. (aunque esto es suficiente para concluir que dicho conjunto no es un espacio vectorial, notemos que ni la suma ni la multiplicación por escalar satisfacen la propiedad clausurativa). □

Ejemplo 3. Sea $\mathcal{F}[0;1]$ el conjunto de todas funciones de valor real definidas en el intervalo $[0;1]$. Si definimos suma y multiplicación por escalar como lo hacemos habitualmente en cálculo: dadas f y g, dos funciones de valor real definidas en $[0;1]$, tenemos que

$$(f+g)(x) = f(x) + g(x) \quad \text{y} \quad (\lambda f)(x) = \lambda[f(x)].$$

En otras palabras, la suma de dos funciones f y g es la función $(f+g)$ tal que el valor que toma en x es la suma de los valores que toman f y g en x. De manera similar, el producto por escalar de λ por f es la función λf, cuyo valor en x es λ veces el valor de f en x.

De la definición de las operaciones, es claro que los Axiomas 1 y 6 se satisfacen. Tomando **0**, el elemento nulo, como la función cuyas imágenes son todas 0, se satisface el Axioma 4 y tomando $(-1)f$ como el opuesto de f se satisface el Axioma 5. La verificación de los demás axiomas es inmediata y se deja como ejercicio para el lector. Así, $\mathcal{F}[0;1]$ es un espacio vectorial. □

Es fácil ver que al tomar cualquier intervalo $[a;b]$ en lugar del intervalo $[0;1]$ en el ejemplo anterior, tenemos otro gran número de ejemplos de espacio vectorial.

Ejemplo 4. Otro ejemplo sencillo de un conjunto que **no** es espacio vectorial real es Z, los números enteros, con la suma y multiplicación por escalar habitual, ya que el Axioma 6 no se satisface: al multiplicar un entero por un número real como $\sqrt{2}$, el resultado no es un número entero. □

Ejemplo 5. Sea $\mathcal{H}$ el hiperplano de R^4 cuya ecuación es $2x_1 - x_2 + 3x_4 = 0$. Verifiquemos que $\mathcal{H}$, con la suma y la multiplicación por escalar definidas en R^4, es un espacio vectorial.

Si $P = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix}$ y $Q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix}$ están en $\mathcal{H}$, tenemos que

$$2p_1 - p_2 + 3p_4 = 0 \quad \text{y} \quad 2q_1 - q_2 + 3q_4 = 0.$$

Los Axiomas 1 y 6 se satisfacen, ya que $(P+Q) = \begin{pmatrix} p_1+q_1 \\ p_2+q_2 \\ p_3+q_3 \\ p_4+q_4 \end{pmatrix}$ y $\lambda P = \begin{pmatrix} \lambda p_1 \\ \lambda p_2 \\ \lambda p_3 \\ \lambda p_4 \end{pmatrix}$ están en $\mathcal{H}$, puesto que

$$2(p_1+q_1) - (p_2+q_2) + 3(p_4+q_4) = (2p_1 - p_2 + 3p_4) + (2q_1 - q_2 + 3q_4) = 0 + 0 = 0$$

y

$$2(\lambda p_1) - (\lambda p_2) + 3(\lambda p_4) = \lambda(2p_1 - p_2 + 3p_4) = \lambda 0 = 0.$$

Es claro que $\mathbf{0} \in \mathcal{H}$: $2 \cdot 0 - 0 + 3 \cdot 0 = 0$ y que $-P \in \mathcal{H}$: $2(-p_1) - (-p_2) + 3(-p_4) = -(2p_1 - p_2 + 3p_4) = -0 = 0$. Así que se satisfacen los Axiomas 4 y 5. Los demás axiomas se satisfacen claramente, ya que se satisfacen para los elementos de R^4 (en particular, para los elementos de $\mathcal{H}$). Así que $\mathcal{H}$ es un espacio vectorial. □

Este último ejemplo es un caso especial, ya que $\mathcal{H}$ es un espacio vectorial contenido en otro espacio vectorial (R^4). Los espacios vectoriales como $\mathcal{H}$ tienen un nombre especial, por su importancia tanto práctica como teórica.

4.3. Subespacio Vectorial

Definición 2 [*Subespacio vectorial*]. Sea V un espacio vectorial y sea H un subconjunto de V. Si H es espacio vectorial con los mismos escalares y las mismas operaciones que V, decimos que H es un *subespacio* de V.

Una de las ventajas del concepto de subespacio es que para demostrar que un subconjunto de un espacio vectorial es en si mismo un espacio vectorial, no es necesario mostrar que se satisfacen los 10 axiomas, como vimos en el Ejemplo 5; basta con verificar solo dos de los 10 axiomas de la definición, como lo establece el siguiente teorema.

Teorema 1 [*Caracterización de subespacio*].
Sea V un espacio vectorial y H un subconjunto no vacío de V. H es un subespacio vectorial de V, si y solo si, los elementos de H satisfacen las propiedades clausurativas para la suma y el producto por escalar (Axiomas 1 y 6).

Demostración: Supongamos que H es un subespacio vectorial de V, entonces H es un espacio vectorial, lo que implica que H satisface los 10 axiomas de la definición, en particular, el 1 y el 6.
Supongamos ahora que H satisface los Axiomas 1 y 6. Teniendo en cuenta que los elementos de H también son elementos de V, H satisface también los Axiomas 2, 3, 7, 8, 9 y 10 (en otras palabras, H hereda estas propiedades de V). Nos queda por verificar únicamente los Axiomas 4 y 5. Sea $\mathbf{u} \in H$. Por el Axioma 6, $\lambda\mathbf{u} \in H$ para todo λ, en particular para $\lambda = 0$ y para $\lambda = -1$. Asi que el elemento nulo, $0\mathbf{u} = \mathbf{0}$, y el opuesto de $\mathbf{u}$, $(-1)\mathbf{u} = -\mathbf{u}$, están en H. □

Así, en el Ejemplo 5, para probar que $\mathcal{H}$ es un subespacio vectorial de R^4, bastaba con verificar sólo los Axiomas 1 y 6.

Ejemplo 6. Demuestre que $\mathcal{D}_n$, el conjunto de matrices diagonales de tamaño $n \times n$, es un espacio vectorial.

Sabiendo que $\mathcal{M}_{n\times n}$, el conjunto de las matrices de tamaño $n \times n$, es un espacio vectorial, que $\mathcal{D}_n$ es no vacío y que $\mathcal{D}_n \subset \mathcal{M}_{n\times n}$, por el teorema anterior, basta con verificar los Axiomas 1 y 6. Pero claramente, al sumar dos matrices diagonales y al multiplicar una matriz diagonal por un escalar, obtenemos de nuevo matrices diagonales. Por el Teorema 1, $\mathcal{D}_n$ es un subespacio vectorial de $\mathcal{M}_{n\times n}$ y por lo tanto, un espacio vectorial. □

Ejemplo 7. Sea $\mathcal{P}_2$ el conjunto de polinomios de grado menor o igual a 2. Demostremos que

$$H = \{p(x) = a_0 + a_1x + a_2x^2 : a_0 = 0,\ a_1, a_2 \in R\}.$$

es un espacio vectorial.

Sabiendo que $\mathcal{P}_2$ es un espacio vectorial y que $H \subset \mathcal{P}_2$, sean $p(x) = a_0 + a_1x + a_2x^2$ y $q(x) = b_0 + b_1x + b_2x^2$ dos polinomios de H. Entonces $a_0 = 0$ y $b_0 = 0$, y por tanto, $a_0 + b_0 = 0$ y $\lambda a_0 = 0$.

Así que

$$p(x) + q(x) = (a_0 + a_1x + a_2x^2) + (b_0 + b_1x + b_2x^2) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$$

y

$$\lambda p(x) = \lambda(a_0 + a_1x + a_2x^2) = (\lambda a_0) + (\lambda a_1)x + (\lambda a_2)x^2$$

son polinomios de H; por el Teorema 1, H es un subespacio vectorial de $\mathcal{P}_2$ y por lo tanto, un espacio vectorial. □

Ejemplo 8. Veamos que $\mathcal{K} = \left\{(0, y, 0, 1 + y)^T : y \in R\right\}$ no es un subespacio vectorial de R^4.

Basta con mostrar que uno de los 10 axiomas de la definición de espacio vectorial no se cumple. Por ejemplo, el vector nulo $(0, 0, 0, 0)^T$ no se encuentra en $\mathcal{K}$ (falla el Axioma 4). □

Ejemplo 9. Demostremos que si $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ son vectores de R^n, entonces $G = Gen\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ es un espacio vectorial.

Veamos que tanto la suma como la multiplicación por escalar son cerradas en G. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores de G y λ un número real. Por la definición de conjunto generado, existen escalares α_i y β_i, para $i = 1,\ 2, \ldots\ n$ tales que

$$\mathbf{u} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_k\mathbf{u}_k \quad \text{y} \quad \mathbf{v} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_k\mathbf{u}_k,$$

por lo tanto,

$$\mathbf{u}+\mathbf{v} = (\alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k) + (\beta_1\mathbf{u}_1 + \cdots + \beta_k\mathbf{u}_k) = (\alpha_1+\beta_1)\mathbf{u}_1 + \cdots + (\alpha_k+\beta_k)\mathbf{u}_k$$

y

$$\lambda\mathbf{u} = \lambda(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_k\mathbf{u}_k) = (\lambda\alpha_1)\mathbf{u}_1 + (\lambda\alpha_2)\mathbf{u}_2 + \cdots + (\lambda\alpha_k)\mathbf{u}_k;$$

es decir, $\mathbf{u}+\mathbf{v} \in G$ y $\lambda\mathbf{u} \in G$. Por el Teorema 1, G es un subespacio de R^n y por tanto, un espacio vectorial.□

Ejemplo 10. Verifiquemos que $H = \left\{ \begin{pmatrix} a \\ a-2b \end{pmatrix} : a, b \in R \right\}$ es un subespacio vectorial de R^2.

Observemos que $\begin{pmatrix} a \\ a-2b \end{pmatrix} = a\begin{pmatrix} 1 \\ 1 \end{pmatrix} + b\begin{pmatrix} 0 \\ -2 \end{pmatrix}$; por lo tanto, $H = Gen\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix} \right\}$; y por el resultado del ejemplo anterior, H es un espacio vectorial. □

4.4. Conceptos Básicos

Como dijimos en la introducción, en esta sección, extenderemos los conceptos básicos de combinación lineal, conjunto generador, conjunto generado e independencia lineal que se estudiaron para vectores de R^n en el Capítulo 2, a los vectores abstractos (elementos de un espacio vectorial).

En un espacio vectorial, dado un conjunto de elementos del espacio, se puede construir un sin número de elementos del espacio al combinar las operaciones básicas, como lo planteamos en la siguiente definición.

Definición 3 [*Combinación lineal*]. Dados V, un espacio vectorial, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, vectores de V, y $\lambda_1, \lambda_2, \ldots, \lambda_n$, escalares, decimos que

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n$$

es una *combinación lineal* de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. A los escalares $\lambda_1, \lambda_2, \cdots, \lambda_n$ se les llama *coeficientes* de la combinación lineal. Si todos los escalares son cero, diremos que tenemos la *combinación lineal trivial* de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

Ejemplo 11. Dado el espacio vectorial $\mathcal{P}_2$, el polinomio $-3 - 6x + x^2$ es una combinación lineal de los polinomios $1 - 2x + x^2$ y $3 + x^2$, puesto que

$$-3 - 6x + x^2 = 3(1 - 2x + x^2) - 2(3 + x^2).$$

Pero, el polinomio $-2x + x^2$ no es una combinación lineal de los polinomios mencionados, ya que no existen escalares α y β tales que

$$\begin{aligned} -2x + x^2 &= \alpha(1 - 2x + x^2) + \beta(3 + x^2) \\ &= (\alpha + 3\beta) - 2\alpha x + (\alpha + \beta)x^2, \end{aligned}$$

ya que el sistema de ecuaciones lineales[1]

$$\begin{aligned} 1 &= \alpha + \beta \\ -2 &= -2\alpha \\ 0 &= \alpha + 3\beta \end{aligned}$$

es inconsistente. □

[1]Recordemos que dos polinomios son iguales, si y solo si, sus coeficientes correspondientes son iguales, es decir, $p(x) = q(x)$, si y solo si, $a_0 = b_0$, $a_1 = b_1, \ldots, a_n = b_n$, siendo $p(x) = a_0 + a_1x + \ldots, +a_nx_n$ y $q(x) = b_0 + b_1x + \ldots, +b_nx_n$.

Al igual que con los vectores de R^n, el conjunto de todas las combinaciones lineales de un conjunto de vectores tiene un nombre especial.

Definición 4 [*Conjunto generado y conjunto generador*]. Sean V, un espacio vectorial, y $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vectores de V. Definimos *conjunto generado* por $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ al conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, lo que representaremos por

$$Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}.$$

Además, si $H = Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, diremos que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un *conjunto generador* de H.

Ejemplo 12. Verifiquemos que $\{1 - x,\ x^2,\ 1 + x\}$ es un conjunto generador de $\mathcal{P}_2$.
Debemos demostrar que cualquier polinomio de grado menor o igual a 2 es combinación lineal de los polinomios $1 - x,\ x^2,\ 1 + x$. En otras palabras, que dado un polinomio $a_0 + a_1x + a_2x^2$, existen escalares $\lambda_1, \lambda_2, \lambda_3 \in R$, tales que $a_0 + a_1x + a_2x^2 = \lambda_1(1 - x) + \lambda_2x^2 + \lambda_3(1 + x)$ o lo que es equivalente, que $a_0 + a_1x + a_2x^2 = (\lambda_1 + \lambda_3) + (\lambda_3 - \lambda_1)x + \lambda_2x^2$. Por la igualdad entre polinomios, el problema se reduce a determinar si el sistema de ecuaciones lineales

$$\begin{aligned} a_0 &= \lambda_1 + \lambda_3 \\ a_1 &= \lambda_3 - \lambda_1 \\ a_2 &= \lambda_2 \end{aligned}$$

tiene solución para todo $(a_0, a_1, a_2)^T$. Al escalonar la matriz de coeficientes del sistema, obtenemos $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.
Como todas las filas tienen pivote, el sistema anterior siempre tiene solución, de donde podemos concluir que, efectivamente, cualquier polinomio de grado menor o igual a 2 es combinación lineal de los polinomios $1 - x,\ x^2,\ 1 + x$; es decir, conforman un conjunto generador de $\mathcal{P}_2$. □

Ejemplo 13. Determinemos si $M = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \right\}$ es un conjunto generador de $\mathcal{M}_{2\times 2}$, el espacio vectorial de las matrices de tamaño 2×2.

Veamos si cualquier matriz de tamaño 2×2 pertenece a $Gen\ M$. Sea $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}$. Determinemos si existen escalares α, β, γ tales que $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \alpha + \gamma & \beta \\ \beta + \gamma & \alpha \end{pmatrix}$, lo que es equivalente, por la igualdad de matrices, a determinar si el sistema de ecuaciones lineales

$$\begin{aligned} a &= \alpha + \gamma \\ b &= \beta \\ c &= \beta + \gamma \\ d &= \alpha \end{aligned}$$

tiene solución para todo $(a, b, c, d)^T$. Como la matriz de coeficientes de este sistema y una forma escalonada equivalente son, respectivamente,

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

podemos concluir que el sistema no siempre tiene solución y por tanto, no toda matriz de $\mathcal{M}_{2\times 2}$ es una combinación lineal de las matrices de M (verifique que la matriz $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ es una de ellas). Por lo tanto,

M no es un conjunto generador de $\mathcal{M}_{2\times 2}$.

□

Ejemplo 14. Determinemos si $S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$ genera a R^3.

Veamos si cualquier vector de R^3 pertenece a *Gen* S. Sea $(x_1, x_2, x_3)^T \in R^3$ y verifiquemos si existen escalares $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ tales que

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_4 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \\ \lambda_2 + \lambda_3 + 2\lambda_4 \\ \lambda_3 + 3\lambda_4 \end{pmatrix},$$

lo que es equivalente a determinar si el sistema de ecuaciones lineales

$$\begin{aligned} x_1 &= \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \\ x_2 &= \lambda_2 + \lambda_3 + 2\lambda_4 \\ x_3 &= \lambda_3 + 3\lambda_4 \end{aligned}$$

tiene solución para todo $(x_1, x_2, x_3)^T$. Como la matriz de coeficientes de este sistema, que coincide con su forma escalonada, es

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix},$$

podemos concluir, que todo vector de R^3 es una combinación lineal de los vectores de S. Por tanto, S es un conjunto generador de R^3. □

En general, el conjunto generado por cualquier subconjunto no vacío de vectores de un espacio vectorial es un subespacio vectorial (y por tanto, un espacio vectorial); además, es el subespacio más pequeño[2] que contiene el subconjunto de vectores, como lo planteamos en el siguiente teorema.

Teorema 2 [*Subespacio mínimo que contiene un subconjunto*].
Si S es un subconjunto de un espacio vectorial V, entonces *Gen* S es un subespacio de V. Además, si H es otro subespacio de V que contiene a S, entonces H contiene a *Gen* S.

Demostración: Aunque el teorema es válido para situaciones más generales, probaremos el resultado solo para cuando S es un conjunto finito. Probemos primero que *Gen* S es subespacio de V.

Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ y sean $\mathbf{u}_1, \mathbf{u}_2 \in$ *Gen* S, de modo que

$$\begin{aligned} \mathbf{u}_1 &= \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k \\ \mathbf{u}_2 &= \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \ldots + \mu_k\mathbf{v}_k. \end{aligned}$$

Entonces,

$$\begin{aligned} \mathbf{u}_1 + \mathbf{u}_2 &= (\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k) + (\mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \ldots + \mu_k\mathbf{v}_k) \\ &= (\lambda_1 + \mu_1)\mathbf{v}_1 + (\lambda_2 + \mu_2)\mathbf{v}_2 + \ldots + (\lambda_k + \mu_k)\mathbf{v}_k \end{aligned}$$

y

$$\begin{aligned} \alpha\mathbf{u}_1 &= \alpha(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k) \\ &= (\alpha\lambda_1)\mathbf{v}_1 + (\alpha\lambda_2)\mathbf{v}_2 + \ldots + (\alpha\lambda_k)\mathbf{v}_k. \end{aligned}$$

Por tanto, $\mathbf{u}_1 + \mathbf{u}_2$ y $\alpha\mathbf{u}_1$ son combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$; es decir, se encuentran en *Gen* S, por lo cual, podemos concluir que *Gen* S es un subespacio de V.

[2]Más pequeño en el sentido de contenencia de conjuntos.

Ahora, probemos que *Gen* S es el subespacio más pequeño que contiene a S. Sea H un subespacio vectorial de V tal que $S \subseteq H$. Como los elementos de H satisfacen las propiedades clausurativas para la suma y el producto por escalar, cualquier combinación lineal de los elementos de S pertenecen a H; es decir, *Gen* $S \subseteq H$. □

Uno de los conceptos más importantes del álgebra lineal es el de dependencia lineal de un conjunto de elementos de un espacio vectorial, el cual está íntimamente relacionado con la forma de escribir el vector $\mathbf{0}$ como combinación lineal de estos elementos, lo cual expresamos en la siguiente definición.

Definición 5 [*Conjunto de vectores linealmente dependiente*]. Sea V un espacio vectorial. Diremos que el conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subset V$ es *linealmente dependiente*(*l.d.*), si existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$, **no todos cero**, tales que

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n = \mathbf{0}.$$

En caso contrario, diremos que el conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es linealmente independiente (*l.i.*); es decir, si siempre que $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n = \mathbf{0}$, se tiene que $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$; en otras palabras, un conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ es *linealmente independiente*, si la única combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ igual a cero es la combinación lineal trivial.

Ejemplo 15. Determinemos si el conjunto de polinomios $S = \{1 - x + x^2, 1 + x - x^2, 1 + x + x^2\}$ es *l.d.* Encontremos los escalares λ_1, λ_2 y λ_3 tales que

$$\begin{aligned} \lambda_1(1 - x + x^2) + \lambda_2(1 + x - x^2) + \lambda_3(1 + x + x^2) &= \mathbf{0}, \\ (\lambda_1 + \lambda_2 + \lambda_3) + (-\lambda_1 + \lambda_2 + \lambda_3)x + (\lambda_1 - \lambda_2 + \lambda_3)x^2 &= 0 \end{aligned}$$

o lo que es equivalente, resolvamos el sistema de ecuaciones lineales

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ -\lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ \lambda_1 - \lambda_2 + \lambda_3 &= 0, \end{aligned} \tag{4.1}$$

cuya matriz de coeficientes y una forma escalonada equivalente son

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix},$$

respectivamente, de donde concluimos que no hay variables libres en el sistema (4.1); es decir, que el sistema homogéneo (4.1) tiene solución única (el vector nulo) y por tanto, S es un conjunto de vectores *l.i.* □

Ejemplo 16. Determinemos si $T = \left\{1 - x,\ 2x - x^2,\ -1 + 2x^2,\ 1 + x + x^2\right\}$ es un conjunto de vectores *l.i.*

Veamos si existen únicos escalares $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, tales que

$$\begin{aligned} \alpha_1(1 - x) + \alpha_2(2x - x^2) + \alpha_3(-1 + 2x^2) + \alpha_4(1 + x + x^2) &= 0, \\ (\alpha_1 - \alpha_3 + \alpha_4) + (-\alpha_1 + 2\alpha_2 + \alpha_4)x + (-\alpha_2 + 2\alpha_3 + \alpha_4)x^2 &= 0, \end{aligned}$$

lo que es equivalente a verificar si existe solución única para el sistema de ecuaciones lineales

$$\begin{aligned} \alpha_1 - \alpha_3 + \alpha_4 &= 0 \\ -\alpha_1 + 2\alpha_2 + \alpha_4 &= 0 \\ -\alpha_2 + 2\alpha_3 + \alpha_4 &= 0, \end{aligned} \tag{4.2}$$

cuya matriz de coeficientes y una forma escalonada equivalente son

$$\begin{pmatrix} 1 & 0 & -1 & 1 \\ -1 & 2 & 0 & 1 \\ 0 & -1 & 2 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 3/2 & 2 \end{pmatrix},$$

respectivamente. La segunda matriz indica que en el sistema (4.2) α_4 es una variable libre; es decir, que dicho sistema tiene infinitas soluciones y por tanto, T no es un conjunto de vectores *l.i.* □

Los conjuntos linealmente dependientes y los conjuntos linealmente independientes tiene propiedades bien interesantes; algunas de ellas los caracterizan, como lo planteamos en el siguiente teorema.

Teorema 3 [*Propiedades de independencia y dependencia lineal*].
Dado un espacio vectorial V y S un subconjunto de V.

1. Si S contiene el vector $\mathbf{0}$, S es un conjunto linealmente dependiente.
2. Si S consta de **dos** elementos, S es linealmente dependiente, si y solo si, uno de ellos es un múltiplo por escalar del otro[3].
3. Si S contiene un subconjunto linealmente dependiente, S es linealmente dependiente.
4. Si S es linealmente independiente, cualquier subconjunto de S es también linealmente independiente.

Demostración: Nuevamente, aunque los Resultados 1, 3 y 4 son válidos en general, demostremos los dos últimos resultados cuando S es un conjunto finito.

1. Dado que $\lambda\mathbf{0} = \mathbf{0}$ para todo λ, el resultado es inmediato.
2. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2\}$. Si S es *l.d.*, entonces existen escalares λ_1 y λ_2 tales que $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 = \mathbf{0}$, donde al menos uno de los dos escalares es diferente de cero. Sin pérdida de generalidad, supongamos que $\lambda_1 \neq 0$, así que $\mathbf{v}_1 = -\dfrac{\lambda_2}{\lambda_1}\mathbf{v}_2$. Ahora, si $\mathbf{v}_1 = \mu\mathbf{v}_2$, entonces podemos obtener una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ igual $\mathbf{0}$, $\mathbf{v}_1 - \mu\mathbf{v}_2 = \mathbf{0}$, que no es la trivial; por tanto, S es un conjunto de vectores *l.d.*
3. Sea $T = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ un subconjunto de $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$. Si T es *l.d.*, existen $\lambda_1, \lambda_2, \ldots, \lambda_k$ no todos 0, tales que $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k = \mathbf{0}$. Asi que $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k + 0\mathbf{v}_{k+1} + \ldots + 0\mathbf{v}_n = \mathbf{0}$ y por tanto S es *l.d.*
4. Sea $T = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ un subconjunto de $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$. Si S es *l.i.* entonces $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k + \lambda_{k+1}\mathbf{v}_{k+1} + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$ implica que $\lambda_1 = \lambda_2 = \ldots = \lambda_k = \lambda_{k+1} = \ldots = \lambda_n = 0$. En particular, si $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k = \mathbf{0}$, entonces $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k + 0\mathbf{v}_{k+1} + \ldots + 0\mathbf{v}_n = \mathbf{0}$ lo que implica que $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$; por lo tanto T es un conjunto *l.i.*. □

Establezcamos ahora algunos resultados que relacionan los conceptos de conjunto generador, conjunto generado y conjunto de vectores linealmente independiente.

Teorema 4 [*Propiedades de un conjunto linealmente independiente*].
Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto de vectores *l.i.* de un espacio vectorial V.

1. Si $\mathbf{v} \in Gen\ S$, entonces $\mathbf{v}$ se escribe de manera única como combinación lineal de los vectores de S. En otras palabras, si
$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n \qquad \text{y} \qquad \mathbf{v} = \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \ldots + \mu_n\mathbf{v}_n,$$
entonces
$$\lambda_1 = \mu_1,\ \lambda_2 = \mu_2,\ \ldots,\ \lambda_n = \mu_n.$$
2. Si $\mathbf{v} \notin Gen\ S$, entonces el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}\}$ es *l.i.*

[3]Observemos que este resultado no es válido si S consta de 3 o más elementos; Por su simplicidad, se enuncia este resultado pero es un caso particular del Teorema 5.

Demostración:

1. Supongamos que existen dos combinaciones lineales de $\mathbf{v} \in Gen\ S$, es decir, que existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ y $\mu_1, \mu_2, \ldots, \mu_n$ tales que

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n \qquad \text{y} \qquad \mathbf{v} = \mu_1\mathbf{v}_1 + \ldots + \mu_n\mathbf{v}_n.$$

Restando estas dos ecuaciones, tenemos

$$\begin{aligned} \mathbf{0} &= (\lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n) - (\mu_1\mathbf{v}_1 + \ldots + \mu_n\mathbf{v}_n) \\ &= (\lambda_1 - \mu_1)\mathbf{v}_1 + \ldots + (\lambda_n - \mu_n)\mathbf{v}_n. \end{aligned}$$

Como S es un conjunto de vectores *l.i.*, entonces

$$\lambda_1 - \mu_1 = \lambda_2 - \mu_2 = \ldots = \lambda_n - \mu_n = 0,$$

es decir,

$$\lambda_1 = \mu_1,\ \lambda_2 = \mu_2,\ \ldots,\ \lambda_n = \mu_n.$$

2. Si $\mathbf{v} \notin Gen\ S$, entonces $\mathbf{v} \neq \mathbf{0}$, puesto que $Gen\ S$ es un espacio vectorial. Razonando por contradicción, supongamos que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}\}$ es un conjunto de vectores *l.d.* Por definición, existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n, \lambda_0$, no todos iguales a cero, tales que

$$\mathbf{0} = \lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n + \lambda_0\mathbf{v},$$

con $\lambda_0 \neq 0$, puesto que si $\lambda_0 = 0$, implicaría que todos los demás escalares también serían cero ya que S es un conjunto *l.i.* Entonces, podemos despejar $\mathbf{v}$,

$$\mathbf{v} = -\frac{\lambda_1}{\lambda_0}\mathbf{v}_1 - \ldots - \frac{\lambda_n}{\lambda_0}\mathbf{v}_n\ ,$$

lo que significa que $\mathbf{v} \in Gen\ S$, lo cual contradice nuestra hipótesis inicial. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}\}$ es un conjunto de vectores *l.i.* □

Observemos que cuando tenemos un conjunto de vectores *l.d.*, existe por lo menos un elemento del conjunto que "sobra", en el sentido que dicho elemento es una combinación lineal de los demás y que el espacio generado por el conjunto, con este elemento o sin él, es el mismo. En el siguiente teorema, resaltamos estas dos características de los conjuntos de vectores *l.d.*, que por lo anterior, le dan significado a su nombre.

Teorema 5 [*Propiedades de un conjunto linealmente dependiente*].
Sean V un espacio vectorial y $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto de vectores de V. Entonces, las siguientes proposiciones son equivalentes.

1. El conjunto S es *l.d.*
2. Existe $\mathbf{v}_i \in S$ tal que $\mathbf{v}_i$ es combinación lineal del resto de vectores de S.
3. Existe $\mathbf{v}_i \in S$ tal que $Gen\ S = Gen\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1},\ \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$.

Demostración: Demostremos primero que la Proposición 1 implica la Proposición 2. Supongamos que S es un conjunto de vectores *l.d.* Entonces existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$, no todos iguales a cero, tales que

$$\mathbf{0} = \lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n.$$

Sea $\lambda_i \neq 0$, entonces

$$\begin{aligned} \mathbf{v}_i &= -\tfrac{\lambda_1}{\lambda_i}\mathbf{v}_1 - \ldots - \tfrac{\lambda_{i-1}}{\lambda_i}\mathbf{v}_{i-1} - \tfrac{\lambda_{i+1}}{\lambda_i}\mathbf{v}_{i+1} - \ldots - \tfrac{\lambda_n}{\lambda_i}\mathbf{v}_n \\ &= \beta_1\mathbf{v}_1 + \ldots + \beta_{i-1}\mathbf{v}_{i-1} + \beta_{i+1}\mathbf{v}_{i+1} + \ldots + \beta_n\mathbf{v}_n\ , \end{aligned} \tag{4.3}$$

es decir, $\mathbf{v}_i$ es una combinación lineal del resto de vectores de S.

Veamos ahora que la Proposición 2 implica la Proposición 3. Supongamos que existe $\mathbf{v}_i \in S$, tal que $\mathbf{v}_i$ es combinación lineal del resto de vectores de S. Sin pérdida de generalidad, podemos suponer que el vector que es combinación lineal del resto de vectores de S es $\mathbf{v}_1$; es decir, podemos suponer que

$$\mathbf{v}_1 = \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n. \tag{4.4}$$

Llamemos $S_1 = \{\mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Es obvio que $Gen(S_1) \subseteq Gen(S)$. Demostremos que $Gen(S) \subseteq Gen(S_1)$, para concluir que $Gen(S_1) = Gen(S)$. Sea $\mathbf{v} \in Gen(S)$, por lo tanto, existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$, tales que

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n.$$

Reemplazando (4.4) en la última expresión, tenemos que

$$\begin{aligned} \mathbf{v} &= \lambda_1(\alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n) + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n \\ &= (\lambda_1\alpha_2 + \lambda_2)\mathbf{v}_2 + \ldots + (\lambda_1\alpha_n + \lambda_n)\mathbf{v}_n \ , \end{aligned}$$

de donde concluimos que $\mathbf{v} \in Gen(S_1)$.

Por último, veamos que la Proposición 3 implica la Proposición 1. Supongamos que existe $\mathbf{v}_i \in S$, tal que $Gen\ S = Gen\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1},\ \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$. Sin pérdida de generalidad, podemos suponer que $Gen\ S = Gen\ S_1$, donde $S_1 = \{\mathbf{v}_2, \ldots \mathbf{v}_n\}$. Como $\mathbf{v}_1 \in S$, entonces $\mathbf{v}_1 \in Gen(S) = Gen(S_1)$, así que existen $\alpha_2, \ \ldots, \ \alpha_n$ tales que $\mathbf{v}_1 = \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$. De esta igualdad, tenemos que

$$\mathbf{v}_1 - \alpha_2\mathbf{v}_2 - \ldots - \alpha_n\mathbf{v}_n = \mathbf{0},$$

la cual es una combinación lineal no trivial igual a $\mathbf{0}$, de donde concluimos que S es *l.d.* □

Ejemplo 17. Determinemos si el conjunto de matrices $\left\{ \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & -3 \end{pmatrix} \right\}$ es linealmente dependiente.
Puesto que

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ -1 & -3 \end{pmatrix},$$

por el teorema anterior, concluimos que el conjunto de matrices dado es linealmente dependiente. □

4.5. Bases y Dimensión

En esta sección, presentaremos un concepto fundamental en el análisis de la estructura de espacio vectorial. Se trata del concepto de base, el cual define un conjunto de vectores que permite describir eficientemente[4] el espacio vectorial y sus propiedades.

Definición 6 [*Base*]. Sea $\mathcal{B}$ un subconjunto no vacío del espacio vectorial V. Diremos que $\mathcal{B}$ es una *base* de V, si y sólo si, el conjunto $\mathcal{B}$ satisface las siguientes condiciones

1. $\mathcal{B}$ es un conjunto linealmente independiente.
2. $\mathcal{B}$ es un conjunto generador de V.

Ejemplo 18. Demostremos que $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es una base de R^n.

Cualquier combinación lineal de $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ igual a un vector dado, nos conduce a un sistema de ecuaciones lineales, cuya matriz de coeficientes es la matriz idéntica de tamaño $n \times n$, de donde podemos concluir que

[4] La propiedad de eficiencia de una base de un espacio vectorial está relacionada con el hecho que es un conjunto óptimo (suficiente y necesario) para construir todo el espacio, expresando todos y cada uno de sus elementos de forma única.

la única combinación lineal igual al vector $\mathbf{0}$ es la trivial ($\mathcal{B}$ es un conjunto *l.i.*) y que cualquier vector se puede expresar como una única combinación lineal de ellos($\mathcal{B}$ genera a R^n). Por tanto, $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es una base de R^n. A esta base la llamamos *base canónica* o *base estándar* de R^n. □

Ejemplo 19. Demostremos que $\mathcal{B} = \{1, x, \ldots, x^n\}$ es una base de $\mathcal{P}_n$.

Es fácil ver que al establecer una ecuación en $\mathcal{P}_n$, en la que un polinomio de grado menor o igual a n, $p(x) = a_0 + a_1x + \ldots + a_nx^n$, es una combinación lineal de $1, x, \ldots, x^n$ con los escalares de la combinación lineal como incógnitas; es decir,

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n = \lambda_0 1 + \lambda_1 x + \ldots + \lambda_n x^n,$$

por la igualdad de polinomios, obtenemos un sistema de ecuaciones lineales cuya matriz de coeficientes es la matriz idéntica de tamaño $(n+1) \times (n+1)$. Así, podemos concluir que cualquier polinomio de grado menor o igual a n se puede expresar de manera única como combinación lineal de los polinomios $1, x, \ldots, x^n$, incluyendo el polinomio $\mathbf{0}$. Por tanto, $\{1, x, \ldots, x^n\}$ es un conjunto generador de $\mathcal{P}_n$ y además es *l.i.*, así que es una base de $\mathcal{P}_n$. A esta base la llamamos *base canónica o base estándar de* $\mathcal{P}_n$. □

Ejemplo 20. Demostremos que $\mathcal{B} = \{E_{11}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, \ldots, E_{m1}, \ldots, E_{mn}\}$, donde E_{ij} es la matriz $m \times n$ con todas sus componentes 0, a excepción de la componente (i, j), que es 1, es una base de $\mathcal{M}_{m\times n}$, el espacio vectorial de las matrices de tamaño $m \times n$.

De nuevo, tenemos que al establecer una ecuación en $\mathcal{M}_{m\times n}$, en la que una matriz $m \times n$ es una combinación lineal de $E_{11}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, \ldots, E_{m1}, \ldots, E_{mn}$, con los escalares de la combinación lineal como incógnitas, obtenemos un sistema de ecuaciones lineales cuya matriz de coeficientes es la matriz idéntica de tamaño $mn \times mn$. Así, podemos concluir que cualquier matriz de tamaño $m \times n$ se puede expresar de manera única como combinación lineal de las matrices $E_{11}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, \ldots, E_{m1}, \ldots, E_{mn}$, incluyendo la matriz nula. Por tanto, $\{E_{11}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, \ldots, E_{m1}, \ldots, E_{mn}\}$ es un conjunto generador de $\mathcal{M}_{m\times n}$ y además es *l.i.*, así que es una base de $\mathcal{M}_{m\times n}$. A esta base la llamamos *base canónica* o *base estándar* de $\mathcal{M}_{m\times n}$. □

Ejemplo 21. Demostremos que $\mathcal{B} = \{1 - x,\ 1 + x,\ x^2\}$ es una base de $\mathcal{P}_2$.

Dado un polinomio arbitrario de $\mathcal{P}_2$, $a_0 + a_1x + a_2x^2$, al establecer la ecuación

$$\begin{aligned} a_0 + a_1x + a_2x^2 &= \lambda_1(1 - x) + \lambda_2(1 + x) + \lambda_3x^2 \\ &= (\lambda_1 + \lambda_2) + (-\lambda_1 + \lambda_2)x + \lambda_3x^2, \end{aligned}$$

por la igualdad de polinomios, obtenemos el sistema de ecuaciones lineales

$$\begin{aligned} a_0 &= \lambda_1 + \lambda_2 \\ a_1 &= -\lambda_1 + \lambda_2 \\ a_2 &= \lambda_3. \end{aligned} \tag{4.5}$$

Como la matriz de coeficientes de este sistema y una forma escalonada equivalente son, respectivamente,

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

podemos concluir que el sistema (4.5) tiene solución (única) para todo $(a_0, a_1, a_2)^T$, así que $\{1 - x, 1 + x, x^2\}$ es un conjunto generador de $\mathcal{P}_2$ y, además, es *l.i.*, ya que el sistema homogéneo asociado a (4.5) también tiene solución única, la trivial; por lo tanto, es una base de $\mathcal{P}_2$. □

Ejemplo 22. Determinemos si $\mathcal{B} = \{1 - x + 3x^2,\ 1 - x,\ x^2\}$ es una base de $\mathcal{P}_2$.

Dado que $1 - x + 3x^2 = (1 - x) + 3(x^2)$, por el Teorema 5, $\mathcal{B}$ es un conjunto *l.d.*, por tanto, $\mathcal{B}$ no es una base. □

Ejemplo 23. Determinemos si $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \right\}$ es una base de $\mathcal{M}_{2\times 2}$.

Sea $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matriz arbitraria de $\mathcal{M}_{2\times 2}$ y veamos si existen escalares α, β, γ tales que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \beta \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} \alpha + \gamma & -2\beta + 2\gamma \\ \beta + 3\gamma & -\alpha + \gamma \end{pmatrix},$$

lo que es equivalente a determinar si el sistema de ecuaciones lineales

$$\begin{aligned} a &= \alpha + \gamma \\ b &= -2\beta + 2\gamma \\ c &= \beta + 3\gamma \\ d &= -\alpha + \gamma \end{aligned} \tag{4.6}$$

tiene solución para todo $(a, b, c, d)^T$. Como la matriz de coeficientes de este sistema y una forma escalonada equivalente son, respectivamente,

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 2 \\ 0 & 1 & 3 \\ -1 & 0 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix},$$

el sistema (4.6) no tiene solución para toda matriz de $\mathcal{M}_{2\times 2}$ (por ejemplo, para la matriz $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, el sistema (4.6) no tiene solución), de donde concluimos que S no genera a $\mathcal{M}_{2\times 2}$, por lo tanto, S no es una base de este espacio. □

Como lo pudimos apreciar en los ejemplos anteriores, cada espacio vectorial tiene al menos una base. Aunque la demostración de este resultado está fuera de nuestro alcance, lo enunciamos por su importancia dentro del desarrollo teórico y práctico del Álgebra Lineal.

Teorema 6 [*Existencia de una base*].
Todo espacio vectorial, excepto el espacio vectorial trivial $V = \{\mathbf{0}\}$, tiene al menos una base.

En los ejemplos, también pudimos observar que al plantear las ecuaciones que nos permiten concluir que un conjunto de vectores es una base, siempre llegamos a un sistema de ecuaciones lineales con solución única, lo que se puede enunciar como el siguiente teorema.

Teorema 7 [*Caracterización de una base*].
Un subconjunto $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de un espacio vectorial V es una base de V, si y solo si, para cada vector $\mathbf{v}$ de V existen escalares únicos $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n.$$

Demostración: Si $\mathcal{B}$ es una base de V, $Gen\ \mathcal{B} = V$, $\mathcal{B}$ es un conjunto *l.i.* y por el Teorema 4, todo vector de V se puede escribir como una combinación lineal única de los vectores de $\mathcal{B}$.

Recíprocamente, si suponemos que todo vector de V es combinación lineal de los vectores de $\mathcal{B}$, entonces $\mathcal{B}$ genera a V. Y, si las combinaciones lineales son únicas, la combinación lineal igual a $\mathbf{0}$ es la trivial, por tanto $\mathcal{B}$ es *l.i.* Así que si todo vector se escribe de manera única como combinación lineal de los vectores de $\mathcal{B}$, $\mathcal{B}$ es una base de V. □

Teniendo en cuenta los Ejemplos 19 y 21, observamos que el espacio vectorial $\mathcal{P}_2$ tiene al menos dos bases con algo en común: el número de elementos, lo cual no es casual. Para probar que todas las bases de un mismo Espacio Vectorial tienen el mismo número de vectores, primero observemos que las bases deben contener

a lo más igual número de elementos que un conjunto generador, ya que no existe un conjunto linealmente independiente con más elementos que un conjunto generador, lo que demostramos en el siguiente teorema.

Teorema 8 [*Propiedad maximal de un conjunto l.i.*]
Si el espacio vectorial V tiene un conjunto generador de n elementos, entonces cualquier subconjunto de V con más de n elementos es *l.d.* En otras palabras, cualquier conjunto de vectores *l.i.* de V tiene a lo sumo el número de vectores de un conjunto generador.

Demostración: Sean $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto generador de V y $T = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subset V$ un conjunto de vectores *l.i.* Veamos que $m \leq n$. Construyamos el conjunto $S_1 = \{\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. S_1 es *l.d.*, ya que S es un conjunto generador de V y $\mathbf{u}_1 \in V$, entonces existe un $\mathbf{v}_i$ que es combinación lineal del resto de vectores de S_1. Sin pérdida de generalidad (reenumerando, si es necesario), supongamos que $i = 1$ y definamos $S_1' = S_1 - \{\mathbf{v}_1\} = \{\mathbf{u}_1, \mathbf{v}_2,\ \mathbf{v}_3,\ \ldots,\ \mathbf{v}_n\}$. Por el Teorema 5, tenemos que $V = Gen\ S_1 = Gen\ S_1'$. Construyamos $S_2 = \{\mathbf{u}_2, \mathbf{u}_1, \mathbf{v}_2,\ \mathbf{v}_3, \ldots, \mathbf{v}_n\}$. S_2 es *l.d.*, ya que S_1' es un conjunto generador de V y $\mathbf{u}_2 \in V$, entonces existe un $\mathbf{v}_i$ que es combinación lineal del resto de vectores de S_2. Sin pérdida de generalidad (reenumerando, si es necesario), supongamos que $i = 2$ y definamos $S_2' = S_2 - \{\mathbf{v}_2\} = \{\mathbf{u}_2, \mathbf{u}_1,\ \mathbf{v}_3,\ \ldots,\ \mathbf{v}_n\}$. Por el Teorema 5, tenemos que $V = Gen\ S_2 = Gen\ S_2'$.

Repitiendo la misma argumentación, podemos construir el conjunto $S_m = \{\mathbf{u}_m,\ \ldots,\ \mathbf{u}_1, \mathbf{v}_m,\ ,\ldots, \mathbf{v}_n\}$ el cual es *l.d.*, para lo cual se requiere que $m \leq n$ porque si $m > n$, tendríamos que $S_{n-1} = \{\mathbf{u}_{n-1},\ \ldots,\ \mathbf{u}_1\}$ es un conjunto de vectores *l.d.*, lo cual es una contradicción, puesto que $T = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ es un conjunto de vectores *l.i.* (ver Resultado 4 del Teorema 3) □

Ahora, podemos probar la observación que hicimos referente a los Ejemplos 19 y 21, la cual es muy importante, ya que nos permitirá definir la dimensión de un espacio vectorial.

Teorema 9 [*Característica común de las bases*].
Todas las bases de un mismo espacio vectorial tienen igual número de elementos.

Demostración: Sean $\mathcal{B}_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}_2 = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ bases de V. Asi que $\mathcal{B}_1$ es un conjunto generador de V y $\mathcal{B}_2$ es *l.i.*, entonces, por el teorema anterior, $m \leq n$. Pero de otro lado, $\mathcal{B}_2$ es conjunto generador de V y $\mathcal{B}_1$ es *l.i.*, entonces, por el teorema anterior, $n \leq m$. Por tanto, $n = m$. □

Definición 7 [*Dimensión*]. Si las bases de un espacio vectorial V tienen n elementos, diremos que la dimensión de V es n, lo que expresaremos como

$$dim(V) = n.$$

En caso que $V = \{\mathbf{0}\}$, espacio vectorial que no tiene base, por conveniencia, diremos que el espacio vectorial tiene dimensión 0, y en caso que un espacio vectorial no tenga una base finita, diremos que es de dimensión infinita.

Así, de los Ejemplos 18, 19 y 20, tenemos que $dim(R^n) = n$, $dim(\mathcal{P}_n) = n + 1$ y $dim(\mathcal{M}_{m\times n}) = mn$. De otro lado, $\mathcal{P}_\infty$, el espacio vectorial de todos los polinomios, es de dimensión infinita. Veamos otros ejemplos menos triviales.

Ejemplo 24. Determinemos la dimensión de $V = Gen\{1 - x, 1 + x, x\}$.
Puesto que $x = \frac{1}{2}(1 + x) - \frac{1}{2}(1 - x)$, tenemos que $V = Gen\{1 - x, 1 + x\} = Gen\{1 - x, 1 + x, x\}$. Ahora, puesto que $1+x$ y $1-x$ no son múltiplos por escalar entre si, el conjunto formado por ellos es *l.i.*, por tanto, $\{1 - x, 1 + x\}$ es una base de V, lo que implica que la dimensión de V es 2. □

Ejemplo 25. Demostremos que cualquier plano en R^n que contenga el origen tiene dimensión 2.
Sea $\mathcal{P}$ un plano de R^n que contiene al origen. Si $\mathbf{u}_1$ y $\mathbf{u}_2$ son los vectores directores de $\mathcal{P}$, entonces $\mathcal{P} = Gen\{\mathbf{u}_1, \mathbf{u}_2\}$ y por tanto, $\mathcal{P}$ es subespacio de R^n (ver Ejemplo 9). Pero, como por definición de plano, $\mathbf{u}_1$ y $\mathbf{u}_2$ no son paralelos, $\mathbf{u}_1$ y $\mathbf{u}_2$ forman un conjunto *l.i.*, asi que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base de $\mathcal{P}$ y por tanto, su dimensión es 2. □

Conociendo la dimensión de un espacio vectorial, se puede determinar el mínimo número de elementos de

un conjunto generador y el máximo de un conjunto *l.i.*, como lo expresa el siguiente teorema.

Teorema 10 [*Propiedad maximal y minimal de un subconjunto de vectores*].
Sea V un espacio vectorial de dimensión n y sea S un subconjunto de V con m elementos.

1. Si S es *l.i.*, entonces $m \leq n$.

2. Si S genera a V, entonces $m \geq n$.

Demostración:

1. Sea $\mathcal{B}$ una base de V. Entonces $\mathcal{B}$ tiene n elementos y genera a V. Por el Teorema 8, cualquier conjunto de vectores *l.i.* tiene a lo sumo n elementos, así que si S es *l.i.*, $m \leq n$.

2. Sea $\mathcal{B}$ una base de V. Entonces $\mathcal{B}$ tiene n elementos y es un conjunto *l.i.* Así que si S genera a V, por el Teorema 8, $m \geq n$. □

Ejemplo 26. Sea S un conjunto de k matrices de tamaño $m \times n$, entonces si S es *l.i.*, $k \leq mn$ y si S genera a $\mathcal{M}_{m\times n}$, entonces $k \geq mn$, ya que $dim(\mathcal{M}_{m\times n}) = mn$. □

Observemos que este teorema nos da condiciones necesarias, muy fáciles de verificar, para que un conjunto de vectores sea generador o linealmente independiente cuando conocemos la dimensión del espacio vectorial, lo cual nos permite determinar cuando un conjunto de vectores no es generador o no es *l.i.*, simplemente comparando el número de vectores del conjunto con la dimensión del respectivo espacio vectorial. Así por ejemplo, sabiendo que $dim(\mathcal{M}_{2\times 2}) = 4$, el conjunto M del Ejemplo 13 no genera a $\mathcal{M}_{2\times 2}$, por tener solo 3 matrices. Igualmente, sabiendo que $dim(R^3) = 3$, el conjunto S del Ejemplo 14 no es *l.i.*, por tener 4 vectores. Lo mismo sucede con el conjunto T del Ejemplo 16, el cual tiene 4 polinomios y $dim(\mathcal{P}_2) = 3$.

Con este teorema, también tenemos que, conociendo n, la dimensión de un espacio vectorial, basta que un conjunto de n elementos satisfaga una de las dos condiciones de la definición de base, para concluir que el conjunto es una base del espacio vectorial.

Teorema 11 [*Propiedad maximal y minimal de una base*].
Sea V un espacio vectorial de dimensión n y sea S un subconjunto de V con n elementos.

1. Si S es *l.i.*, entonces S es una base de V.

2. Si S genera a V, entonces S es una base de V.

Demostración:

1. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto de vectores *l.i.* de V. Si S no genera a V, existe $\mathbf{u}$ en V que no se encuentra en $Gen\ S$ y por tanto, el conjunto $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de $n+1$ elementos es *l.i.*, lo que nos lleva a una contradicción, ya que, por el Teorema 10, tendríamos que $(n+1) \leq n$. De donde concluimos que S genera a V y por tanto, es una base de V.

2. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto generador de V. Si S es *l.d.*, por el Teorema 5, existe un vector $\mathbf{v}_i$ que es combinación lineal de los otros vectores de S, asi que si $S_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$, $Gen\ S = Gen\ S_1$, pero S_1 contiene $n-1$ elementos, lo que contradice la parte 2 del Teorema 10. Así que S es *l.i.*, lo que implica que S es una base de V. □

Este teorema nos muestra otra ventaja importante de conocer la dimensión de un espacio vectorial cuando queremos determinar si un conjunto dado es una base de él. Así por ejemplo, conociendo que $dim(\mathcal{P}_2) = 3$, el resultado obtenido en el Ejemplo 12 nos permite afirmar que el conjunto $\{1 - x,\ x^2,\ 1 + x\}$ es una base de $\mathcal{P}_2$.

Para enfatizar la propiedad de eficiencia de una base, tenemos que, dado un conjunto generador pero *l.d.* de un espacio vectorial V, podemos quitarle elementos adecuadamente hasta reducirlo a un conjunto *l.i.*

y por tanto, en base de V. Y si tenemos un conjunto de vectores *l.i.* de V, podemos agregarle vectores adecuadamente hasta convertirlo en generador de V y por tanto, en base de V, como lo demostramos en el siguiente teorema.

Teorema 12 [*Construcción de una base a partir de un subconjunto de vectores*].
Sea V un espacio vectorial de dimensión n y S un subconjunto de V con m elementos.

1. Si S es *l.i.*, podemos encontrar un conjunto T que contiene a S y es una base de V.

2. Si $V = Gen\ S$, entonces podemos encontrar un conjunto T contenido en S que es una base de V.

Demostración:

1. Supongamos que S es *l.i.* Si S genera a V, $T = S$ es una base de V. Si S no genera a V, existe un elemento $\mathbf{u}_1 \in V$ tal que $\mathbf{u}_1 \notin Gen\ S$. Construyamos $S_1 = S \cup \{\mathbf{u}_1\}$; si S_1 genera a V, $T = S_1$ es una base de V ya que, por el Teorema 4, es *l.i.*; sino, existe un elemento $\mathbf{u}_2 \in V$ tal que $\mathbf{u}_2 \notin Gen\ S_1$. Construyamos a $S_2 = S_1 \cup \{\mathbf{u}_2\}$, si S_2 genera a V, $T = S_2$ es una base de V ya que, por el Teorema 4, es *l.i.*; sino, podemos continuar este procedimiento hasta encontrar un conjunto S_k que contiene a S, es *l.i.* y genera a V. Por tanto, podemos encontrar un conjunto $T = S_k$ (por el Teorema 11, $k = n-m$) que contiene a S y es una base de V.

2. Si S es *l.i.*, entonces S es una base de V. Si S es *l.d.*, entonces existe un vector $\mathbf{u}_1$ de S que es combinación lineal de los otros vectores de S de tal manera que si $S_1 = S - \{\mathbf{u}_1\}$, $Gen\ S = Gen\ S_1$. Si S_1 es *l.i.*, entonces S_1 es una base de V. Si S_1 es *l.d.*, entonces existe un vector $\mathbf{u}_2$ de S_1 que es combinación lineal de los otros vectores de S_1 de tal manera que si $S_2 = S_1 - \{\mathbf{u}_2\}$, $Gen\ S_2 = Gen\ S_1 = Gen\ S$ y continuamos este procedimiento hasta obtener un conjunto $T = S_k$ (por el Teorema 11, $k = m - n$) que es *l.i.*, tal que $Gen\ S_k = Gen\ S$. Este conjunto $T = S_k$ es una base de V. □

Ejemplo 27. Encontremos una base de $\mathcal{P}_2$ que contenga a $S = \{1 - x^2, 1 + 2x\}$.

Como S es *l.i.*, basta encontrar un vector de $\mathcal{P}_2$ que no esté en $Gen\ S$. Así, tendríamos un conjunto de 3 vectores *l.i.*, que, por el Teorema 11, sería base de V.

Sabiendo que $\mathcal{B} = \{1, x, x^2\}$ es una base de $\mathcal{P}_2$, tomemos un vector de $\mathcal{B}$, por ejemplo 1 y veamos si está en $Gen\ S$. Es fácil ver que $1 \notin Gen\ S$, pues no existen escalares λ_1, λ_2 tales que $\lambda_1(1 - x^2) + \lambda_2(1 + 2x) = 1$. Así que $S_1 = \{1 - x^2, 1 + 2x, 1\}$ tiene 3 elementos y es *l.i.*, lo que implica que S_1 es una base de $\mathcal{P}_2$. □

Para terminar esta sección, demostraremos un resultado que seguramente nos parece natural, pero que es importante resaltar: la dimensión de un subespacio vectorial es a lo sumo la dimensión del espacio vectorial que lo contiene y el único subespacio con la misma dimensión del espacio vectorial que lo contiene es el mismo espacio vectorial.

Teorema 13 [*Máxima dimensión de un subespacio*].
Sean V un espacio vectorial de dimensión n y W un subespacio vectorial de V. Entonces,

1. $dim W \leq n$.

2. Si $dim W = n$, entonces $W = V$.

Demostración:

1. Sea $\mathcal{B}$ una base de W. Como $\mathcal{B}$ es un conjunto *l.i.* de V, por el Teorema 10, $\mathcal{B}$ tiene a lo sumo n vectores y por tanto, $dim W \leq n$.

2. Si $dim W = n$ y $\mathcal{B}$ es una base de W, entonces $W = Gen\ \mathcal{B}$ y $\mathcal{B}$ tiene n vectores *l.i.* de W y por tanto de V. Por el Teorema 11, $\mathcal{B}$ es una base de V, lo que implica que $Gen\ \mathcal{B} = V$ y por tanto, $W = V$. □

4.6. Coordenadas Respecto a una Base Ordenada.

Por el Teorema 7, dada una base de un espacio vectorial, cada vector del espacio se escribe de manera única como una combinación lineal de los elementos de la base. Así que los coeficientes de la combinación lineal identifican al vector, como lo establece la siguiente definición.

Definición 8 [*Vector de coordenadas*]. Sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base (ordenada) del espacio vectorial V. Como para todo $\mathbf{v} \in V$, existen escalares únicos $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{v},$$

al vector de R^n

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

lo llamamos *vector de coordenadas* de $\mathbf{v}$ respecto a la base ordenada $\mathcal{B}$.

Claramente, si cambiamos el orden de los vectores de una base, el vector de coordenadas puede cambiar, asi como cuando cambiamos la base.

Ejemplo 28. Demostremos que el vector de coordenadas de un vector $\mathbf{v}$ de R^n, respecto a la base canónica, coincide con el vector $\mathbf{v}$.

Sea $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ un vector de R^n. Así que $\mathbf{v} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$. Por tanto, $[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, donde $\mathcal{B}$ es la base canónica, por lo que concluimos que $\mathbf{v} = [\mathbf{v}]_{\mathcal{B}}$. □

Ejemplo 29. Calculemos el vector de coordenadas de $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ respecto a la base ordenada

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right\}.$$

Tenemos que encontrar $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ tales que

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 - \lambda_2 & \lambda_2 + \lambda_3 \\ \lambda_1 + \lambda_4 & \lambda_3 + \lambda_4 \end{pmatrix}.$$

Al resolver el sistema de ecuaciones lineales resultante, encontramos que $\lambda_1 = 1, \ \lambda_2 = 0, \ \lambda_3 = 0$ y $\lambda_4 = -1$; es decir,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 1 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} - 1 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Por lo tanto,

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}.$$

□

Ejemplo 30. Calculemos $p(x)$, sabiendo que $[p(x)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -1 \\ 3 \\ -2 \end{pmatrix}$ y que $\mathcal{B} = \left\{1 - x, 2x + x^2, x^2 - x^3, x^3 - 1\right\}$ es una base de $\mathcal{P}_3$.

Por definición de vector de coordenadas,

$$p(x) = 1(1-x) - 1(2x+x^2) + 3(x^2-x^3) - 2(x^3-1) = 3 - 3x + 2x^2 - 5x^3.$$

□

Ejemplo 31. Observemos que al cambiar el orden de los elementos de $\mathcal{B}$ en el ejemplo anterior, las coordenadas de $p(x)$, respecto a la nueva base ordenada, cambian.

Sea $\mathcal{B}' = \{2x+x^2, x^2-x^3, x^3-1, 1-x\}$. Como

$$p(x) = -1(2x+x^2) + 3(x^2-x^3) - 2(x^3-1) + 1(1-x),$$

entonces $[p(x)]_{\mathcal{B}'} = \begin{pmatrix} -1 \\ 3 \\ -2 \\ 1 \end{pmatrix}$ y por tanto, $[p(x)]_{\mathcal{B}'} \neq [p(x)]_{\mathcal{B}}$. □

Otro resultado curioso y supremamente útil es que los coeficientes de una combinación lineal de vectores de un vector dado se conservan al cambiar los vectores por los vectores de coordenadas de estos, respecto a una base dada, lo cual consignamos en el siguiente teorema.

Teorema 14 [*Conservación de los coeficientes de una combinación lineal*].
Sean $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ vectores de un espacio vectorial V y $\mathcal{B}$ una base de V. Entonces $\mathbf{u} = \lambda_1\mathbf{u}_1 + \ldots + \lambda_k\mathbf{u}_k$, si y solo si, $[\mathbf{u}]_{\mathcal{B}} = \lambda_1 [\mathbf{u}_1]_{\mathcal{B}} + \ldots + \lambda_k [\mathbf{u}_k]_{\mathcal{B}}$.

Demostración: Sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base del espacio vectorial V y supongamos que

$$[\mathbf{u}_i]_{\mathcal{B}} = \begin{pmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{ni} \end{pmatrix} \qquad \text{para} \quad i = 1, 2, \ldots, k;$$

es decir,

$$\mathbf{u}_i = u_{1i}\mathbf{v}_1 + u_{2i}\mathbf{v}_2 + \ldots + u_{ni}\mathbf{v}_n.$$

Por consiguiente, si

$$\mathbf{u} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \ldots + \lambda_k\mathbf{u}_k,$$

entonces,

$$\begin{aligned} \mathbf{u} &= \lambda_1(u_{11}\mathbf{v}_1 + u_{21}\mathbf{v}_2 + \ldots + u_{n1}\mathbf{v}_n) + \ldots + \lambda_k(u_{1k}\mathbf{v}_1 + u_{2k}\mathbf{v}_2 + \ldots + u_{nk}\mathbf{v}_n) \\ &= (\lambda_1 u_{11} + \lambda_2 u_{12} + \ldots + \lambda_k u_{1k})\mathbf{v}_1 + \ldots + (\lambda_1 u_{n1} + \lambda_2 u_{n2} + \ldots + \lambda_k u_{nk})\mathbf{v}_n. \end{aligned}$$

De aqui, que

$$\begin{aligned} [\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 u_{11} + \lambda_2 u_{12} + \ldots + \lambda_k u_{1k} \\ \vdots \\ \lambda_1 u_{n1} + \lambda_2 u_{n2} + \ldots + \lambda_k u_{nk} \end{pmatrix} &= \lambda_1 \begin{pmatrix} u_{11} \\ \vdots \\ u_{n1} \end{pmatrix} + \ldots + \lambda_k \begin{pmatrix} u_{1k} \\ \vdots \\ u_{nk} \end{pmatrix} \\ &= \lambda_1 [\mathbf{u}_1]_{\mathcal{B}} + \ldots + \lambda_k [\mathbf{u}_k]_{\mathcal{B}}, \end{aligned}$$

terminando la demostración de una implicación. Puesto que cada paso es realmente una equivalencia, podemos devolvernos y tener la otra implicación. □

Ejemplo 32. Verifiquemos el resultado del teorema, con la combinación lineal de

$$2x^2 - 2x + 9 = \mathbf{2}(x^2 - 2x + 1) + \mathbf{3}(x+2) \mathbf{-1}(x-1),$$

para las bases $\mathcal{B} = \left\{1, x, x^2\right\}$ y $\mathcal{B}' = \left\{1, 1+x, 1+x+x^2\right\}$ de $\mathcal{P}_2$.

No es difícil ver que

$$[2x^2-2x+9]_{\mathcal{B}} = \begin{pmatrix} 9 \\ -2 \\ 2 \end{pmatrix},\ [x^2-2x+1]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},\ [x+2]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix},\ [x-1]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

y

$$\begin{pmatrix} 9 \\ -2 \\ 2 \end{pmatrix} = \mathbf{2}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \mathbf{3}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \textbf{-1}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Con cálculos similares, tenemos que

$$[2x^2-2x+9]_{\mathcal{B}'} = \begin{pmatrix} 11 \\ -4 \\ 2 \end{pmatrix},\ [x^2-2x+1]_{\mathcal{B}'} = \begin{pmatrix} 3 \\ -3 \\ 1 \end{pmatrix},\ [x+2]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},\ [x-1]_{\mathcal{B}'} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

y

$$\begin{pmatrix} 11 \\ -4 \\ 2 \end{pmatrix} = \mathbf{2}\begin{pmatrix} 3 \\ -3 \\ 1 \end{pmatrix} + \mathbf{3}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \textbf{-1}\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}.$$

□

El resultado del Teorema 14 permite plantear un problema de un espacio vectorial de dimensión n que involucre combinaciones lineales como un problema similar en término de vectores de R^n, como lo muestra el siguiente ejemplo.

Ejemplo 33. Determinemos si $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right\}$ es un conjunto de matrices *l.i.*

Sea

$$\mathcal{B} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

la base canónica de $\mathcal{M}_{2\times 2}$. Por consiguiente,

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ \left[\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \text{ y } \left[\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix}.$$

Es fácil ver que el conjunto de vectores $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix}\right\}$ es *l.i.* Es decir, que la única combinación lineal de ellos igual al vector nulo es la trivial, así que la única combinación lineal de las matrices del conjunto

$$\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right\}$$

es la trivial, de donde concluimos que $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right\}$ es un conjunto *l.i.* □

Los vectores de coordenadas de un vector dado, respecto a dos bases diferentes, están relacionados entre si mediante el producto por una matriz, como lo establece el siguiente teorema.

Teorema 15 [*Matriz de transición*].
Sean $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ bases de un espacio vectorial V y sea P la matriz de tamaño $n \times n$, cuyas columnas son los vectores de coordenadas de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ respecto a la base $\mathcal{B}'$; es decir,

$$P = \left[[\mathbf{v}_1]_{\mathcal{B}'} \ [\mathbf{v}_2]_{\mathcal{B}'} \ \cdots \ [\mathbf{v}_n]_{\mathcal{B}'} \right].$$

Entonces, para cada vector $\mathbf{v} \in V$,

$$[\mathbf{v}]_{\mathcal{B}'} = P \, [\mathbf{v}]_{\mathcal{B}}$$

A la matriz P la llamamos *matriz de transición*[5] de $\mathcal{B}$ a $\mathcal{B}'$.

Demostración: Sea $\mathbf{v} \in V$. Como $\mathcal{B}$ genera a V, existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots \lambda_n \mathbf{v}_n.$$

En otras palabras,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Por el Teorema 14,

$$[\mathbf{v}]_{\mathcal{B}'} = \lambda_1 \, [\mathbf{v}_1]_{\mathcal{B}'} + \lambda_2 \, [\mathbf{v}_2]_{\mathcal{B}'} + \ldots + \lambda_n \, [\mathbf{v}_n]_{\mathcal{B}'},$$

por tanto,

$$[\mathbf{v}]_{\mathcal{B}'} = P \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = P \, [\mathbf{v}]_{\mathcal{B}} \ ,$$

donde

$$P = \left[[\mathbf{v}_1]_{\mathcal{B}'} \ [\mathbf{v}_2]_{\mathcal{B}'} \ \cdots \ [\mathbf{v}_n]_{\mathcal{B}'} \right].$$

□

Algunos resultados interesantes sobre las matrices de transición hacen referencia a que estas matrices son únicas e invertibles y que la inversa de la matriz de transición de una base a otra es la inversa de la matriz de transición de la segunda base a la primera, lo cual lo planteamos en el siguiente teorema.

Teorema 16 [*Unicidad e invertibilidad de la matriz de transición*].
Sean $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ bases de un espacio vectorial V y sea P la matriz de transición de la base $\mathcal{B}$ a la base $\mathcal{B}'$. Es decir,

$$P = \left[[\mathbf{v}_1]_{\mathcal{B}'} \ [\mathbf{v}_2]_{\mathcal{B}'} \ \cdots \ [\mathbf{v}_n]_{\mathcal{B}'} \right].$$

1. Si existe otra matriz P' tal que $[\mathbf{v}]_{\mathcal{B}'} = P' \, [\mathbf{v}]_{\mathcal{B}}$, entonces $P = P'$.

2. La matriz de transición P es invertible y su inversa es la matriz de transición de la base $\mathcal{B}'$ a la base $\mathcal{B}$.

Demostración:

1. Sabemos que para cada vector $\mathbf{v} \in V$, $[\mathbf{v}]_{\mathcal{B}'} = P \, [\mathbf{v}]_{\mathcal{B}}$ y que en particular,

$$[\mathbf{v}_i]_{\mathcal{B}'} = P \, [\mathbf{v}_i]_{\mathcal{B}} = P\mathbf{e}_i \qquad \text{para cada } i = 1, 2, \ldots, n. \tag{4.7}$$

[5]Esta matriz también recibe el nombre de *matriz cambio de base de $\mathcal{B}$ a $\mathcal{B}'$*

Supongamos que existe P' tal que para cada vector $\mathbf{v} \in V$, $[\mathbf{v}]_{\mathcal{B}'} = P' [\mathbf{v}]_{\mathcal{B}}$. Así, en particular,

$$[\mathbf{v}_i]_{\mathcal{B}'} = P' [\mathbf{v}_i]_{\mathcal{B}} = P'\mathbf{e}_i \qquad \text{para cada } i = 1, 2, \ldots, n. \tag{4.8}$$

Por lo tanto, de (4.7) y (4.8),

$$P\mathbf{e}_i = P'\mathbf{e}_i \quad \text{para cada } i = 1, 2, \ldots, n.$$

Como $P\mathbf{e}_i$ y $P'\mathbf{e}_i$ son las i-ésimas columnas de P y P', respectivamente, concluimos que $P = P'$.

2. Sabiendo que $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de V, por el Teorema 14, $\{[\mathbf{v}_1]_{\mathcal{B}'}, [\mathbf{v}_2]_{\mathcal{B}'}, \ldots, [\mathbf{v}_n]_{\mathcal{B}'}\}$ es una base de R^n; por lo tanto, $P = [[\mathbf{v}_1]_{\mathcal{B}'} \ [\mathbf{v}_2]_{\mathcal{B}'} \ \cdots \ [\mathbf{v}_n]_{\mathcal{B}'}]$ es una matriz invertible.
De otro lado, al multiplicar por P^{-1} la igualdad $[\mathbf{v}]_{\mathcal{B}'} = P [\mathbf{v}]_{\mathcal{B}}$, obtenemos

$$P^{-1} [\mathbf{v}]_{\mathcal{B}'} = [\mathbf{v}]_{\mathcal{B}},$$

para todo $\mathbf{v} \in V$. Por la unicidad de la matriz de transición, podemos concluir que P^{-1} es la matriz de transición de la base $\mathcal{B}'$ a la base $\mathcal{B}$. □

Ejemplo 34. Sean $\mathcal{B} = \{1, 1+x, 1+x+x^2\}$ y $\mathcal{B}'$ bases de $\mathcal{P}_2$, donde $\mathcal{B}'$ es la base canónica. Encontremos las matrices de transición de una base a la otra.

Puesto que $\mathcal{B}'$ es la base canónica, $[1]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $[1+x]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ y $[1+x+x^2]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, por consiguiente, la matriz de transición de $\mathcal{B}$ a $\mathcal{B}'$ es $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Calculemos ahora la matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$, para lo cual tenemos que resolver los sistemas que resultan de cada una de las siguientes ecuaciones

$$\begin{aligned} 1 &= a_{11}(1) + a_{21}(1+x) + a_{31}(1+x+x^2) = (a_{11}+a_{21}+a_{31}) + (a_{21}+a_{31})x + a_{31}x^2 \\ x &= a_{12}(1) + a_{22}(1+x) + a_{32}(1+x+x^2) = (a_{12}+a_{22}+a_{32}) + (a_{21}+a_{32})x + a_{32}x^2 \\ x^2 &= a_{13}(1) + a_{23}(1+x) + a_{33}(1+x+x^2) = (a_{13}+a_{23}+a_{33}) + (a_{23}+a_{33})x + a_{33}x^2 \end{aligned}$$

Aplicando los algoritmos de eliminación de Gauss y de sustitución hacía atrás a los sistemas de ecuaciones lineales

$$\begin{aligned} a_{11}+a_{21}+a_{31} &= 1 \\ a_{21}+a_{31} &= 0 \\ a_{31} &= 0 \end{aligned}, \quad \begin{aligned} a_{12}+a_{22}+a_{32} &= 0 \\ a_{21}+a_{32} &= 1 \\ a_{32} &= 0 \end{aligned}, \quad \begin{aligned} a_{13}+a_{23}+a_{33} &= 0 \\ a_{23}+a_{33} &= 0 \\ a_{33} &= 1, \end{aligned}$$

tenemos que

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Por tanto, la matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$ es

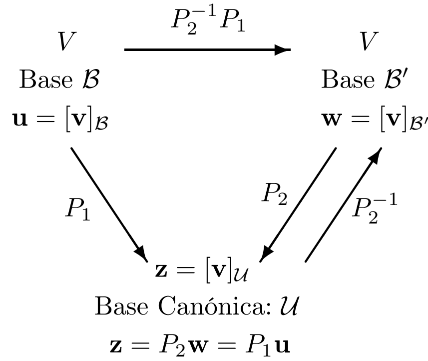
$$P' = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Podemos verificar que $PP' = I$, asi que $P^{-1} = P'$. □

El ejemplo anterior, además de verificar lo establecido en el Teorema 16, nos muestra un hecho importante y que en general es cierto. La matriz de transición de una base a la base canónica es muy fácil de calcular:

basta con encontrar los vectores de coordenadas de la base dada en términos de la base canónica; pero, para hallar la matriz de transición de la base canónica a otra base, necesitamos resolver varios sistemas de ecuaciones o, equivalentemente, calcular la inversa de la matriz de transición anterior, lo cual, además de costoso en términos de cálculos, es innecesario para hacer un cambio de coordenadas. Recordemos que, para calcular el producto de la inversa de una matriz por un vector, conociendo la matriz y, el vector, el cálculo más eficiente consiste en resolver el sistema de ecuaciones cuya matriz aumentada esta conformada por la matriz y el vector.

La observación anterior nos permite inferir que una forma eficiente de expresar (no de calcular) la matriz de transición de una base a otra, diferentes a la base canónica, es mediante el producto de la inversa de la matriz de transición de la segunda base a la base canónica por la matriz de transición de la primeraa base a la base canónica. Esto es, si P_1 es la matriz de transición de la base $\mathcal{B}$ a la base canónica y P_2 es la matriz de transición de la base $\mathcal{B}'$ a la base canónica, entonces $P_2^{-1}P_1$ es la matriz de transición de la base $\mathcal{B}$ a la base $\mathcal{B}'$, como lo ilustra el siguiente gráfico.



La ecuación que aparece en la parte inferior del gráfico anterior, $P_2\mathbf{w} = P_1\mathbf{u}$, la llamamos *ecuación cambio de base* ya que permite hallar las coordenadas de un vector en una base a partir de las coordenadas del mismo vector en la otra base.

4.7. Rango y Nulidad de una Matriz

En el Capítulo 2, definimos dos conjuntos importantes relacionados con una matriz A de tamaño $m \times n$, que resultan ser espacios vectoriales. Ellos son el espacio nulo y el espacio columna:

$$N_A = \{\mathbf{x} \in R^n : A\mathbf{x} = \mathbf{0}\}$$

y

$$C_A = \{\mathbf{b} \in R^m : A\mathbf{x} = \mathbf{b},\ \text{para algún } x \in R^n\},$$

respectivamente.

En efecto, por el Teorema 1, para demostrar que dada una matriz A de tamaño $m \times n$, N_A y C_A son subespacios vectoriales de R^n y R^m, respectivamente, basta con demostrar que se satisfacen la propiedades clausurativas tanto para la suma, como para el producto por escalar (Axiomas 1 y 6 de la definición de espacio vectorial), lo cual hicimos en los Teoremas 3 y 4 del Capítulo 2.

Como N_A y C_A son espacios vectoriales, tienen base y dimensión. Estas dimensiones son características importantes de la matriz A, por lo que reciben nombres especiales, los cuales presentamos en las siguientes definiciones.

Definición 9 [*Nulidad de una matriz*]. Dada una matriz A, definimos $\nu(A)$, la nulidad de A, como la dimensión del espacio nulo de A.

Definición 10 [*Rango de una matriz*]. Dada una matriz A, definimos $\rho(A)$, el rango de A, como la dimensión del espacio columna de A.

Ejemplo 35. Determinemos la nulidad y el rango de las matrices

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{pmatrix} \qquad \text{y} \qquad B = \begin{pmatrix} 1 & -2 & 1 & 1 \\ 2 & 0 & 2 & -2 \\ 0 & 8 & 0 & -8 \end{pmatrix}.$$

Es fácil ver que una forma escalonada equivalente de A es

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -5 \end{pmatrix}$$

y por tanto, el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución única para todo $\mathbf{b}$. Así que el espacio nulo y el espacio columna de A son $\{\mathbf{0}\}$ y R^3, respectivamente; por consiguiente, $\nu(A) = 0$ y $\rho(A) = 3$.

De otro lado, una forma escalonada equivalente del sistema $B\mathbf{x} = \mathbf{c}$ es

$$\left(\begin{array}{cccc|c} 1 & -2 & 1 & 1 & c_1 \\ 0 & 4 & 0 & -4 & c_2 - 2c_1 \\ 0 & 0 & 0 & 0 & c_3 - 2c_2 + 4c_1 \end{array}\right).$$

Al resolver el sistema de ecuaciones resultante, por sustitución hacía atrás, para cuando $\mathbf{c} = \mathbf{0}$, obtenemos que

$$N_B = \left\{ x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} : \ x_3, \ x_4 \in R \right\},$$

así que $\nu(B) = 2$. Ahora, puesto que $B\mathbf{x} = \mathbf{c}$ sólo tiene solución cuando $c_3 - 2c_2 + 4c_1 = 0$, tenemos que

$$C_B = \left\{ \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} : c_3 = -4c_1 + 2c_2, \ c_1, \ c_2 \in R \right\} = \left\{ c_1 \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} : \ c_1, \ c_2 \in R \right\}.$$

Por consiguiente, $\rho(B) = 2$. □

En el ejemplo anterior, del sistema de ecuaciones lineales que se plantea para determinar el núcleo o el rango de una matriz, podemos observar que la nulidad de una matriz coincide con el número de variables libres y que el rango coincide con el número de variables pivotales, de tal manera que la suma de la nulidad y el rango de la matriz es el número total de variables o de columnas de la matriz del sistema, conclusión a la que llegaremos después de demostrar un par de resultados previos.

Primero, demostremos que las columnas correspondientes de dos matrices equivalentes[6] forman conjuntos de vectores que son ambos $l.i.$ o ambos $l.d.$

Teorema 17 [*Matrices equivalentes e independencia lineal*].
Si $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \mathbf{a}_n]$ y $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \mathbf{b}_n]$ son matrices equivalentes, entonces $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}\}$ es $l.i.$, si y solo si, $\{\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \ldots, \mathbf{b}_{i_k}\}$ es $l.i.$, para todo $\{i_1, \ i_2, \ \ldots, \ i_k\} \subseteq \{1, \ 2, \ \ldots, \ n\}$.

Demostración: Como A y B son equivalentes, de la teoría del Capítulo 1, sabemos que las soluciones de los sistemas $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ son las mismas. De otro lado, por definición, $A\mathbf{x}$ es una combinación lineal de los vectores columna de A. Por consiguiente, $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n = \mathbf{0}$, si y solo si, $\lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n = \mathbf{0}$.

[6]Recordemos del Capítulo 1 que dos matrices son equivalentes, si y solo si, una de ellas se puede obtener a partir de la otra mediante operaciones elementales entre filas.

En particular, para cualquier subconjunto $\{i_1, i_2, \ldots, i_k\}$ del conjunto $\{1, 2, \ldots, n\}$, $\lambda_{i_1}\mathbf{a}_{i_1} + \lambda_{i_2}\mathbf{a}_{i_2} + \cdots + \lambda_{i_k}\mathbf{a}_{i_k} = \mathbf{0}$, si y solo si, $\lambda_{i_1}\mathbf{b}_{i_1} + \lambda_{i_2}\mathbf{b}_{i_2} + \cdots + \lambda_{i_k}\mathbf{b}_{i_k} = \mathbf{0}$. Así, que si $\lambda_{i_1}\mathbf{a}_{i_1} + \lambda_{i_2}\mathbf{a}_{i_2} + \cdots + \lambda_{i_k}\mathbf{a}_{i_k} = \mathbf{0}$ implica que $\lambda_{i_1} = \lambda_{i_2} = \cdots = \lambda_{i_k} = 0$, entonces $\lambda_{i_1}\mathbf{b}_{i_1} + \lambda_{i_2}\mathbf{b}_{i_2} + \cdots + \lambda_{i_k}\mathbf{b}_{i_k} = \mathbf{0}$ implica que $\lambda_{i_1} = \lambda_{i_2} = \cdots = \lambda_{i_k} = 0$. Por tanto, $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}\}$ es *l.i.*, si y solo si, $\{\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \ldots, \mathbf{b}_{i_k}\}$ es *l.i.* □

Ahora, presentemos una forma de encontrar una base del espacio columna de una matriz dada.

Teorema 18 [*Base del espacio columna de una matriz*].
Dada una matriz A, las columnas de A correspondientes a las columnas pivotales de una forma escalonada equivalente forman una base de C_A.

Demostración: Sea $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_n]$ una matriz de tamaño $m \times n$ y sea $B = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n]$ una matriz escalonada equivalente, obtenida después de aplicarle el método de eliminación de Gauss. Por el Teorema 17, las columnas $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}\}$ de A forman un conjunto *l.i.*, si y solo si, las columnas $\{\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \ldots, \mathbf{b}_{i_k}\}$ de B son *l.i.* y, utilizando el Teorema 6 del Capítulo 2, si $\{\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \ldots, \mathbf{b}_{i_k}\}$ son las columna pivotales de B, ellas forman un conjunto *l.i.*

Nuevamente, por el Teorema 6 del Capítulo 2, si $\{\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \ldots, \mathbf{b}_{i_k}\}$ son las columnas pivotales de B, entonces $\{\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \ldots, \mathbf{b}_{i_k}, \mathbf{b}_{i_j}\}$, con $j \neq 1, 2, \ldots, k$, es un conjunto *l.d.* y, por el Teorema 17, $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}, \mathbf{a}_{i_j}\}$ es un conjunto *l.d.* Ahora, por el Teorema 5, tenemos que $Gen\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}, \mathbf{a}_{i_j}\} = Gen\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}\}$. Si, al conjunto $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}, \mathbf{a}_{i_j}\}$ le continuamos adjuntando el resto de vectores columna de A, utilizando cada vez el mismo resultado del Teorema 5, tenemos que $Gen\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\} = Gen\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}\}$. Por consiguiente, el conjunto $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}\}$ es *l.i.* y genera el espacio columna de A y, por lo tanto, es una base de C_A. □

De un lado, notemos que la base de C_A está formada por columnas de A y no por columnas de la matriz escalonada equivalente. Y de otro lado, notemos que si bien este último teorema nos garantiza que el conjunto de columnas de A **correspondiente** a las columnas pivotales de una forma escalonada de A es una base de C_A, el teorema **no** está diciendo que sea el único conjunto de columnas que sea base de C_A.

Ejemplo 36. Encontremos una base de $V = Gen\left\{\begin{pmatrix} 1 \\ -1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 5 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 9 \\ 0 \end{pmatrix}\right\}$.

Sea A la matriz cuyas columnas son los vectores del conjunto generador. Es decir, sea $A = \begin{pmatrix} 1 & -1 & 2 & 1 & 3 \\ -1 & 0 & 1 & 2 & 3 \\ 1 & -2 & 5 & 4 & 9 \\ 2 & -1 & 1 & -1 & 0 \end{pmatrix}$

y sea $B = \begin{pmatrix} 1 & -1 & 2 & 1 & 3 \\ 0 & -1 & 3 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ una forma escalonada equivalente a A. Como las primera y segunda columnas de B son sus columnas pivotales, por el teorema anterior, podemos concluir que las primera y segunda columnas de A, $\begin{pmatrix} 1 \\ -1 \\ 1 \\ 2 \end{pmatrix}$ y $\begin{pmatrix} -1 \\ 0 \\ -2 \\ -1 \end{pmatrix}$, forman una base de V. Además, observemos que la primera y segunda columna de B **no** forman una base de C_A y que la primera y cuarta columna de A también forman una base de C_A (Por qué?) □

Ejemplo 37. Encontremos una base de $\mathcal{P}_2$ contenida en $\{1 - x, x - 2x^2, -2 + 2x, 1 - 2x^2, 1 + x - x^2\}$.

Si $\mathcal{B}$ es la base canónica de $\mathcal{P}_2$, entonces $[1-x]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $[x-2x^2]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$, $[-2+2x]_{\mathcal{B}} = \begin{pmatrix} -2 \\ 2 \\ 0 \end{pmatrix}$,

$[1-2x^2]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$ y $[1+x-x^2]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$. Por el Teorema 14, un conjunto de vectores es *l.i.*, si y solo si, sus vectores de coordenadas respecto a una base dada forman un conjunto *l.i.* Asi que basta con identificar, de los vectores $[1-x]_{\mathcal{B}}, [x-2x^2]_{\mathcal{B}}, [-2+2x]_{\mathcal{B}}, [1-2x^2]_{\mathcal{B}}, [1+x-x^2]_{\mathcal{B}}$, cuales forman un conjunto *l.i.*

Al calcular una forma escalonada de la matriz cuyas columnas son los vectores de coordenadas de los polinomios dados, obtenemos $\begin{pmatrix} 1 & 0 & -2 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$. Por lo tanto, podemos concluir que los polinomios correspondientes a las columnas pivotales de esta forma escalonada (primera, segunda y quinta columnas) forman una base de $\mathcal{P}_2$; es decir, $\{1-x,\ x-2x^2,\ 1+x-x^2\}$ es una base de $\mathcal{P}_2$. ¿Existe otro subconjunto del conjunto de polinomios dado que sea una base de $\mathcal{P}_2$? ¿Las primera, segunda y quinta columnas de la matriz escalonada anterior forman una base de $\mathcal{P}_2$? □

Después de estos resultados, estamos en capacidad de demostrar el teorema que habíamos anunciado, la suma de la nulidad y el rango de una matriz es igual al número de columnas de ella.

Teorema 19 [*Relación entre nulidad y rango de una matriz*].
Dada una matriz A de tamaño $m \times n$,

$$\nu(A) + \rho(A) = n.$$

Demostración: Sean $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ una base de N_A, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base de C_A y $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ vectores de R^n tales que $A\mathbf{u}_j = \mathbf{v}_j$, para $j = 1, 2, \ldots, k$. Veamos que $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ es una base de R^n, con lo que quedaría demostrado el teorema.

Para tal efecto, comencemos por demostrar que $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ es un conjunto *l.i.* Sea

$$\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \ldots + \alpha_r\mathbf{w}_r + \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \ldots + \beta_k\mathbf{u}_k = \mathbf{0}$$

Al multiplicar ambos lados de la igualdad anterior por la matriz A y teniendo en cuenta que $A(\alpha_i\mathbf{w}_i) = \alpha_i A\mathbf{w}_i = \mathbf{0}$ y $A(\beta_j\mathbf{u}_j) = \beta_j(A\mathbf{u}_j) = \beta_j\mathbf{v}_j$, obtenemos

$$\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_k\mathbf{v}_k = \mathbf{0}.$$

Por la independencia de $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, concluimos que $\beta_1 = \beta_2 = \ldots = \beta_k = 0$, y por tanto,

$$\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \ldots + \alpha_r\mathbf{w}_r + \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \ldots + \beta_k\mathbf{u}_k = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \ldots + \alpha_r\mathbf{w}_r = \mathbf{0}.$$

Ahora, por la independencia de $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$, tenemos que $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$.

Para demostrar que $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ genera a R^n, tomemos un vector $\mathbf{v}$ de R^n. Como $A\mathbf{v} \in C_A$, existen $\beta_1, \beta_2, \ldots, \beta_k$ escalares tales que $A\mathbf{v} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_k\mathbf{v}_k$. Construyamos los vectores $\mathbf{u} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \ldots + \beta_k\mathbf{u}_k$ y $\mathbf{z} = \mathbf{v} - \mathbf{u}$. Como $A\mathbf{z} = A\mathbf{v} - A\mathbf{u}$ y

$$A\mathbf{u} = \beta_1 A\mathbf{u}_1 + \beta_2 A\mathbf{u}_2 + \ldots + \beta_k A\mathbf{u}_k = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_k\mathbf{v}_k = A\mathbf{v},$$

entonces $\mathbf{z} \in N_A$. Por lo tanto, existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$, tales que $\mathbf{z} = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \ldots + \alpha_r\mathbf{w}_r$. Como $\mathbf{v} = \mathbf{z} + \mathbf{u}$, concluimos que existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_k, \beta_1, \beta_2, \ldots, \beta_k$, tales que

$$\mathbf{v} = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \ldots + \alpha_r\mathbf{w}_r + \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \ldots + \beta_k\mathbf{u}_k.$$

□

Corolario 19.1
Dada una matriz A de tamaño $m \times n$, su nulidad es igual al número de variables libres del sistema $A\mathbf{x} = \mathbf{0}$.

Demostración: Por el Teorema 19, tenemos que $\nu(A) + \rho(A) = n$ y, por el Teorema 18, tenemos que las columnas de A, correspondientes a las columnas pivotales de su forma escalonada, forman una base de C_A, asi que $\rho(A)$ es el número de columnas pivotales. Por consiguiente, $\nu(A)$ es el número de variables no pivotales; es decir, el número de variables libres. □

Ejemplo 38. Verifiquemos el resultado del Teorema 19 usando las matrices del Ejemplo 35.

La matriz A tiene 3 columnas, $\nu(A) = 0$ y $\rho(A) = 3$. La matriz B tiene 4 columnas, $\nu(B) = 1$ y $\rho(B) = 3$. □

Ejemplo 39. Si tenemos un sistema de 15 ecuaciones con 20 incógnitas, qué podemos decir de las soluciones del sistema, si sabemos que el espacio nulo tiene dimensión 5?

Sea A la matriz de coeficientes del sistema. Entonces A tiene 15 filas y 20 columnas. Puesto que $\nu(A) = 5$, $\rho(A) = 20 - 5 = 15$. Como, $C_A \subseteq R^{15}$ y $\rho(A) = 15$, por el Teorema 13, $C_A = R^{15}$. Es decir, todo vector de R^{15} es combinación lineal de las columnas de A, así que el sistema siempre tiene solución; además, son infinitas. □

Sabiendo que el conjunto generado por las filas de una matriz es un espacio vectorial, nos preguntamos cómo encontrar una base y/o su dimensión y cuáles son sus relaciones con los espacio columna y nulo de la misma matriz. Comencemos por definir y presentar algunos resultados del espacio generado por los vectores formados por las filas en si y luego estableceremos algunas relaciones con las dimensiones de los espacios columna y nulo.

Definición 11 [*Espacio fila*]. Dada una matriz A de tamaño $m \times n$, definimos F_A, *el espacio fila* de A, como el espacio generado por los vectores formados por las filas de A.

Ejemplo 40. Dada la matriz $A = \begin{pmatrix} 2 & -1 & 0 & 2 \\ 4 & -1 & -3 & 3 \\ 0 & 1 & -3 & -1 \end{pmatrix}$, encontremos el espacio fila de A, una base del mismo y su dimensión.

Por la Definición 11, el espacio fila de A es

$$F_A = \left\{ \mathbf{y} \in R^4, \, \mathbf{y} = \alpha_1 \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 4 \\ -1 \\ -3 \\ 3 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 1 \\ -3 \\ -1 \end{pmatrix}, \, \alpha_i \in R \right\}.$$

Para encontrar una base de F_A, basta determinar cuales de los vectores formados por las filas de A forman un conjunto *l.i.*, para lo cual, escalonamos la matriz cuyas columnas son dichos vectores, es decir, la matriz A^T. Escalonando la matriz

$$A^T = \begin{pmatrix} 2 & 4 & 0 \\ -1 & -1 & 1 \\ 0 & -3 & -3 \\ 2 & 3 & -1 \end{pmatrix} \quad , \quad \text{obtenemos} \quad \begin{pmatrix} 2 & 4 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

de donde podemos afirmar que los vectores $\begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}$ y $\begin{pmatrix} 4 \\ -1 \\ -3 \\ 3 \end{pmatrix}$ (las columnas 1ª y 2ª de A^T) forman un conjunto *l.i.*; por lo tanto, forman una base de F_A y $dim(F_A) = 2$. □

A diferencia de lo que ocurre entre los espacios columna de las matrices respectivas, al aplicar una operación elemental entre filas a una matriz, las filas de la nueva matriz generan el mismo espacio que las de la matriz inicial como veremos en el siguiente teorema.

Teorema 20 [*Base del espacio fila de una matriz*].
Si las matrices A y B son equivalentes, entonces $F_A = F_B$ y si B es una matriz escalonada, el conjunto de vectores formados por las filas de B, diferentes a cero, es *l.i.* y por tanto, es una base de F_B.

Demostración: Basta demostrar que si B es la matriz que se obtiene al aplicarle una operación elemental a A, entonces $F_A = F_B$. Sean $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ los vectores formados por las filas de la matriz A. Si la operación es de tipo permutación, el conjunto de vectores formados por las filas de las dos matrices es el mismo y por tanto, el conjunto generado por las filas de las dos matrices también es el mismo. Si la operación es de tipo escalamiento, $cF_i \longrightarrow F_i$, es fácil ver que si $\mathbf{w} = \lambda_1\mathbf{f}_1 + \ldots + \lambda_i\mathbf{f}_i + \ldots + \lambda_m\mathbf{f}_m$, entonces $\mathbf{w} = \lambda_1\mathbf{f}_1 + \ldots + \frac{\lambda_i}{c}c\mathbf{f}_i + \ldots + \lambda_m\mathbf{f}_m$ y por tanto, el conjunto generado por las filas de las dos matrices es el mismo. Y si la operación es de tipo eliminación, $F_i + cF_j \longrightarrow F_i$, podemos ver que si $\mathbf{w} = \lambda_1\mathbf{f}_1 + \ldots + \lambda_i\mathbf{f}_i + \ldots + \lambda_j\mathbf{f}_j + \ldots + \lambda_m\mathbf{f}_m$, entonces $\mathbf{w} = \lambda_1\mathbf{f}_1 + \ldots + \lambda_i(\mathbf{f}_i + c\mathbf{f}_j) + \ldots + (\lambda_j - \lambda_i c)\mathbf{f}_j + \ldots + \lambda_m\mathbf{f}_m$ y por tanto, el conjunto generado por las filas de las dos matrices también es el mismo.

La demostración de que el conjunto de vectores formados por las filas de la forma escalonada equivalente, diferentes a cero, es *l.i.*, la dejamos como ejercicio para el lector. □

Ejemplo 41. Utilicemos el Teorema 20 para encontrar otra base de F_A, siendo A la matriz dada en el Ejemplo 40.

Al escalonar la matriz A dada en el Ejemplo 40, obtenemos la matriz $U = \begin{pmatrix} 2 & -1 & 0 & 2 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Claramente, los vectores formados por las filas 1^a y 2^a de U forman un conjunto *l.i.*; por lo tanto, forman una base de F_U. Como A y U son matrices equivalentes, por el Teorema 20, $F_U = F_A$, entonces, dicho conjunto es también una base de F_A. □

Observemos que la dimensión del espacio columna de una matriz es el número de columnas pivotales de una forma escalonada equivalente y, en la forma escalonada, cada fila diferente de cero contiene un pivote, lo que implica que las dimensiones de los espacios fila y columna son las mismas, lo que planteamos, sin demostración, en el siguiente teorema.

Teorema 21 [*Relación entre espacio columna y espacio fila de una matriz*].
Dada cualquier matriz A, $dimC_A = dimF_A$.

A lo largo de esta sección, hemos obtenido varios resultados que relacionan los sistemas de ecuaciones lineales con los espacios vectoriales asociados a una matriz, los cuales consignaremos en dos teoremas resumen. La demostración de la mayoría de las implicaciones ya las hemos hecho.

Teorema 22 [*Resumen 1*].
Sea A una matriz $m \times n$. Entonces las siguientes afirmaciones son equivalentes.

1. El número de pivotes de la forma escalonada de A es n.
2. Los vectores columna de A forman un conjunto *l.i.*
3. $\rho(A) = n$
4. $dimC_A = n$.
5. $\rho(A^T) = n$
6. $dimF_A = n$.
7. $\nu(A) = 0$.
8. $N_A = \{\mathbf{0}\}$.
9. El sistema $A\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial.

Teorema 23 [*Resumen 2*].
Sea A una matriz $m \times n$. Entonces las siguientes afirmaciones son equivalentes.

1. El número de pivotes de la forma escalonada de A es m.
2. Los vectores columna de A generan a R^m.
3. $\rho(A) = m$
4. $dim C_A = m$.
5. $\rho(A^T) = m$
6. $dim F_A = m$.
7. Cada fila de la forma escalonada de A contiene un pivote.
8. $\nu(A) = n - m$.
9. El sistema $A\mathbf{x} = \mathbf{b}$ tiene solución para todo $\mathbf{b}$.

Corolario 23.1 Sea A una matriz $n \times n$. Entonces las siguientes afirmaciones son equivalentes.

1. El número de pivotes de la forma escalonada de A es n.
2. Los vectores columna de A generan a R^n.
3. Los vectores columna de A forman un conjunto $l.i.$
4. Los vectores columna de A forman una base de R^n
5. $\rho(A) = n$
6. $dim C_A = n$.
7. $dim F_A = n$.
8. Cada fila de la forma escalonada de A contiene un pivote.
9. $\nu(A) = 0$.
10. El sistema $A\mathbf{x} = \mathbf{b}$ tiene solución única para todo $\mathbf{b}$.
11. La matriz A es invertible.
12. $det A \neq 0$.

4.8. Producto Escalar y Bases Ortonormales en R^n

En el Capítulo 2, para los vectores de R^n, definimos otra operación además de la suma y el producto por escalar: *producto escalar*. Esta operación nos permitió caracterizar los vectores ortogonales: dos vectores $\mathbf{u}$ y $\mathbf{v}$ de R^n son ortogonales, si y solo si, su producto escalar es 0. En general, si en un espacio vectorial definimos una operación que cumpla las propiedades básicas del producto escalar, podemos extender el concepto de ortogonalidad[7].

[7] El producto escalar definido en R^n es sólo un caso particular de lo que conocemos como *producto interno* de un espacio vectorial, y un producto interno en un espacio vectorial V es una función que a cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ asigna un número real $\langle \mathbf{u}, \mathbf{v} \rangle$ y satisface las siguientes propiedades para todo $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ y todo $\alpha \in R$,

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.
2. $\langle (\mathbf{u} + \mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$.
3. $\alpha \langle \mathbf{u}, \mathbf{v} \rangle = \langle \alpha \mathbf{u}, \mathbf{v} \rangle$.
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ ($\langle \mathbf{u}, \mathbf{u} \rangle = 0$, si y solo si, $\mathbf{u} = \mathbf{0}$.)

Sin embargo, en esta sección, nos limitaremos a estudiar las principales propiedades que tiene un conjunto de vectores ortogonales en R^n, ya que por los resultados de la Sección 4.6, tanto el producto escalar como dichas propiedades se pueden extender a cualquier espacio vectorial de dimensión n.[8]

Para empezar, recordemos que el concepto de ortogonalidad está definido entre dos vectores de R^n. Hablaremos de un conjunto de vectores ortogonales, cuando tengamos un conjunto de vectores que dos a dos son ortogonales, como lo planteamos en la siguiente definición.

Definición 12 [*Conjunto ortogonal*]. Decimos que un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ de R^n es ortogonal, si y sólo si, cualquier par de vectores diferentes del conjunto son ortogonales. Es decir, si $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, para todo $i \neq j$.

Ejemplo 42. Veamos que cualquier subconjunto del conjunto de vectores $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de R^n es ortogonal. Es fácil ver que $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ para todo $i \neq j$, ya que las únicas componentes de $\mathbf{e}_i$ y $\mathbf{e}_j$ diferentes de cero son la i-ésima y la j-ésima, respectivamente; por tanto, el producto de componente por componente es cero. □

Los conjuntos ortogonales satisfacen propiedades muy importantes, una de ellas es que son linealmente independientes y su demostración es trivial.

Teorema 24 [*Ortogonalidad e independencia lineal*].
Si $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto de vectores ortogonales de R^n, entonces S es un conjunto de vectores linealmente independiente.

Demostración: Expresemos $\mathbf{0}$ como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$,

$$\mathbf{0} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_k \mathbf{v}_k \tag{4.9}$$

y verifiquemos que $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$.

Al multiplicar escalarmente la ecuación (4.9) por $\mathbf{v}_i$, obtenemos $\mathbf{0} \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i \cdot \mathbf{v}_i$, ya que por ser S un conjunto ortogonal, $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, para $i \neq j$. Así que

$$\lambda_i = \frac{\mathbf{0} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} = 0 \qquad \text{para } i = 1, 2, \ldots, k.$$

□

Siguiendo un razonamiento muy similar al de esta última demostración, podemos calcular fácilmente los coeficientes de la combinación lineal de vectores ortogonales para un vector dado del subespacio generado por ellos, como lo ilustra el siguiente teorema, siendo ésta una de las ventajas más importantes de trabajar con un conjunto de vectores ortogonales.

Teorema 25 [*Ortogonalidad y combinación lineal*].
Dado un conjunto de vectores ortogonales $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ de R^n, si $\mathbf{u} \in Gen\ S$; es decir, si

$$\mathbf{u} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_k \mathbf{v}_k, \tag{4.10}$$

entonces

$$\lambda_i = \frac{\mathbf{u} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \qquad \text{para todo } i = 1, 2, \ldots, k.$$

Demostración: Al multiplicar escalarmente la ecuación (4.10) por $\mathbf{v}_i$, obtenemos $\mathbf{u} \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i \cdot \mathbf{v}_i$, ya que por ser S un conjunto ortogonal, $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, para $i \neq j$. De donde,

$$\lambda_i = \frac{\mathbf{u} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \qquad \text{para } i = 1, 2, \ldots, k.$$

□

[8] Sean $\mathbf{u}, \mathbf{v}$ vectores de un espacio vectorial V y sea $\mathcal{B}$ una base de V; como $[\mathbf{u}]_{\mathcal{B}}, [\mathbf{v}]_{\mathcal{B}} \in R^n$, podemos definir $\langle \mathbf{u}, \mathbf{v} \rangle$, el producto escalar entre $\mathbf{u}$ y $\mathbf{v}$, como $\langle \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_{\mathcal{B}} \cdot [\mathbf{v}]_{\mathcal{B}}$.

Ejemplo 43. Dado el conjunto $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -5 \\ 1 \\ 2 \end{pmatrix} \right\}$, verifiquemos que S es un conjunto ortogonal y escribamos el vector $(-3 \quad -3 \quad 3)^T$ como combinación lineal de los vectores de S.

El conjunto S es ortogonal, ya que

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = 0{+}2{-}2 = 0 \quad , \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 1 \\ 2 \end{pmatrix} = 0{+}2{-}2 = 0 \quad y \quad \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 1 \\ 2 \end{pmatrix} = -5{+}1{+}4 = 0 \,.$$

Por el Teorema 25, basta con calcular

$$\begin{pmatrix} -3 \\ -3 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = -3 - 3 + 6 = 0 \ , \quad \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = 1 + 1 + 4 = 6$$

$$\begin{pmatrix} -3 \\ -3 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = 0 - 6 - 3 = -9 \ , \quad \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = 0 + 4 + 1 = 5$$

$$\begin{pmatrix} -3 \\ -3 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 1 \\ 2 \end{pmatrix} = 15 - 3 + 6 = 18 \quad \text{y} \quad \begin{pmatrix} -5 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 1 \\ 2 \end{pmatrix} = 25 + 1 + 4 = 30$$

y calcular los cocientes $\frac{0}{6} = 0$, $\frac{-9}{5} = -1{,}8$ y $\frac{18}{30} = 0{,}6$, para concluir que

$$\left[\begin{pmatrix} -3 \\ -3 \\ 3 \end{pmatrix} \right]_S = \begin{pmatrix} 0 \\ -9/5 \\ 18/30 \end{pmatrix} = \begin{pmatrix} 0{,}0 \\ -1{,}8 \\ 0{,}6 \end{pmatrix} \ .$$

□

La ortogonalidad de un conjunto de vectores se puede caracterizar en términos del producto de una matriz con su transpuesta, como se expresa en el interesante resultado del siguiente teorema.

Teorema 26 [*Caracterización de un conjunto ortogonal*].
Sea A una matriz cuyas columnas son los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in R^n - \{\mathbf{0}\}$. Entonces, $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ es un conjunto ortogonal de vectores, si y solo si, $A^T A = D = (d_{ij})$ es una matriz diagonal; además, $d_{ii} = \|\mathbf{v}_i\|^2$.

Demostración: Al realizar el producto de las matrices A^T y A, tenemos que si $D = A^T A$,

$$d_{ij} = fil_i(A^T) \cdot col_j(A) = \mathbf{v}_i \cdot \mathbf{v}_j$$

y puesto que $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, si y solo si, $\mathbf{v}_i$ y $\mathbf{v}_j$ son ortogonales para $i \neq j$ y $\mathbf{v}_i \cdot \mathbf{v}_i = \|\mathbf{v}_i\|^2$, entonces

$$d_{ij} = \begin{cases} \|\mathbf{v}_i\|^2 & \text{si } \ i = j \\ 0 & \text{si } \ i \neq j \end{cases} ,$$

si y solo si, $\mathbf{v}_i$ y $\mathbf{v}_j$ son ortogonales. □

Cuando los vectores de un conjunto, además de ser ortogonales, son unitarios, les llamamos de una manera especial.

Definición 13 [*Conjunto ortonormal*]. Diremos que un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ de R^n es ortonormal, si y sólo si, además de ser un conjunto ortogonal, cada uno de los vectores $\mathbf{v}_i$ es unitario. En otras palabras, si y sólo si,

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Ejemplo 44. Cualquier subconjunto no vacío de la base canónica de R^n es un conjunto ortonormal de vectores. □

Ejemplo 45. Verifiquemos que $\mathcal{B} = \left\{ \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}, \begin{pmatrix} 2/\sqrt{14} \\ 1/\sqrt{14} \\ 3/\sqrt{14} \end{pmatrix} \right\}$ es un conjunto ortonormal.

En efecto,

$$\begin{aligned}
\begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix} &= \left(1/\sqrt{3}\right)^2 + \left(1/\sqrt{3}\right)^2 + \left(-1/\sqrt{3}\right)^2 = 1, \\
\begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} 2/\sqrt{14} \\ 1/\sqrt{14} \\ 3/\sqrt{14} \end{pmatrix} &= 2/\sqrt{52} + 1/\sqrt{52} - 3/\sqrt{52} = 0 \\
\begin{pmatrix} 2/\sqrt{14} \\ 1/\sqrt{14} \\ 3/\sqrt{14} \end{pmatrix} \cdot \begin{pmatrix} 2/\sqrt{14} \\ 1/\sqrt{14} \\ 3/\sqrt{14} \end{pmatrix} &= \left(2/\sqrt{14}\right)^2 + \left(1/\sqrt{14}\right)^2 + \left(3/\sqrt{14}\right)^2 = 1
\end{aligned}$$

□

Definición 14 [*Base ortogonal y ortonormal*]. Si $\mathcal{B}$ es una base del subespacio V de R^n y al mismo tiempo es un conjunto ortogonal de vectores, diremos que $\mathcal{B}$ es una *base ortogonal* de V, y si además, $\mathcal{B}$ es un conjunto ortonormal de vectores, diremos que $\mathcal{B}$ es una *base ortonormal* de V.

Ejemplo 46. La base canónica de R^n es una base ortonormal puesto que , además de ser una base de R^n, es un conjunto ortogonal y cada uno de sus elementos tiene norma 1. □

Por el resultado del Teorema 25, una de las principales ventajas de tener una base ortonormal de un subespacio de R^n es que podemos encontrar fácilmente los coeficientes de la combinación lineal de cualquier vector del subespacio en dicha base.

Teorema 27 [*Ortonormalidad y combinación lineal*].
Si $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base ortonormal de un subespacio V de R^n y $\mathbf{u}$ está en V, entonces

$$\mathbf{u} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k, \qquad \text{donde} \qquad \lambda_i = \mathbf{u} \cdot \mathbf{v}_i \quad \text{para } i = 1, 2, \ldots, k.$$

Demostración: Por ser $\mathcal{B}$ una base de V, cualquier vector $\mathbf{u}$ de V, se puede escribir como combinación lineal de los vectores de $\mathcal{B}$, $\mathbf{u} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k$, y, por el Teorema 25, tenemos que $\lambda_i = \frac{\mathbf{u} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}$ para $i = 1, 2, \ldots, k$. Ahora, como los vectores de $\mathcal{B}$ son unitarios, $\lambda_i = \mathbf{u} \cdot \mathbf{v}_i$. □

Ejemplo 47. Sean $V = Gen\ \mathcal{B}$, $\mathbf{u} = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}$ y $\mathbf{w} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$, donde $\mathcal{B}$ es el conjunto dado en el Ejemplo 45. Determinemos si $\mathbf{u}$ y $\mathbf{w}$ pertenecen a V.

Como $\mathcal{B}$ es un conjunto ortonormal, $\mathcal{B}$ es un conjunto *l.i.* y, puesto que por definición de V, $\mathcal{B}$ genera a V, tenemos que $\mathcal{B}$ es una base ortonormal de V. Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ los vectores de $\mathcal{B}$. Para determinar si $\mathbf{u} \in V$, debemos encontrar escalares α_1 y α_2 tales que $\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$. De otro lado, si $\mathbf{u} \in V$, por el Teorema 27,

$\alpha_1 = \mathbf{u}\cdot\mathbf{v}_1 = \frac{3}{\sqrt{3}} = \sqrt{3}$ y $\alpha_2 = \mathbf{u}\cdot\mathbf{v}_2 = \frac{14}{\sqrt{14}} = \sqrt{14}$. En efecto, podemos verificar que $\mathbf{u} = \sqrt{3}\mathbf{v}_1 + \sqrt{14}\mathbf{v}_2$; por lo tanto, $\mathbf{u} \in V$.

En forma similar, para determinar si $\mathbf{w} \in V$, debemos encontrar escalares α_1 y α_2 tales que $\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$. Como $\mathcal{B}$ es una base ortonormal de V, si $\mathbf{w} \in V$, por el Teorema 27, $\alpha_1 = \mathbf{w}\cdot\mathbf{v}_1 = \frac{-3}{\sqrt{3}} = -\sqrt{3}$ y $\alpha_2 = \mathbf{w}\cdot\mathbf{v}_2 = \frac{4}{\sqrt{14}}$. Pero, como $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = -\sqrt{3}\mathbf{v}_1 + \frac{4}{\sqrt{14}}\mathbf{v}_2 = \frac{1}{7}\begin{pmatrix} -3 \\ -5 \\ 13 \end{pmatrix} \neq \mathbf{w}$, concluimos que el vector $\mathbf{w}$ no pertenece al subespacio V. □

4.9. Proyección ortogonal

Comparando el resultado del Teorema 27 con los procedimientos requeridos para calcular el vector de coordenadas de un vector en una base ordenada que no es ortogonal, resulta importante preguntarnos si todos los subespacios vectoriales de R^n tienen bases ortonormales, lo que respondemos afirmativamente al final de esta sección, donde además, ilustramos cómo, a partir de una base del subespacio, podemos encontrar una base ortonormal para el mismo subespacio. Para facilitar la comprensión de dicho procedimiento, generalicemos el concepto de proyección ortogonal de un vector sobre otro, a la proyección ortogonal de un vector sobre un subespacio de R^n, para lo cual introducimos la siguiente definición.

Definición 15. [*Ortogonalidad a un subespacio*]. Diremos que un vector $\mathbf{u} \in R^n$ es ortogonal a un subespacio S de R^n, si y solo si, el vector $\mathbf{u}$ es ortogonal a todos y cada uno de los vectores del subespacio S; es decir,

$$\mathbf{u}\cdot\mathbf{v} = 0 \quad \text{para todo } \mathbf{v} \in S.$$

Ejemplo 48. Sea S el subespacio definido por un hiperplano que pasa por el origen en R^n. Por la definición de hiperplano (Definición 23 del Capítulo 2), su vector normal es ortogonal a S. □

Afortunadamente, no es necesario verificar los infinitos productos escalares que la anterior definición sugiere para verificar la ortogonalidad de un vector a un subespacio; en realidad, es suficiente y necesario que el vector sea ortogonal a un conjunto generador del subespacio, como lo plantea el siguiente teorema.

Teorema 28 [*Caracterización de ortogonalidad de un vector a un subespacio*].
Un vector $\mathbf{u} \in R^n$ es ortogonal al subespacio $S = Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, si y solo si, el vector $\mathbf{u}$ es ortogonal a los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.

Demostración: Si el vector $\mathbf{u}$ es ortogonal a S, por la Definición 15, $\mathbf{u}$ es ortogonal a todos los vectores de S, y en particular, $\mathbf{u}$ es ortogonal a $\mathbf{v}_i$ para $i = 1, 2, \ldots k$. En el otro sentido, si $\mathbf{u}$ es ortogonal a $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, entonces $\mathbf{u}\cdot\mathbf{v}_i = 0$ para $i = 1, 2 \ldots k$. Sea $\mathbf{v} \in S$, entonces existen escalares $\alpha_1, \alpha_2 \ldots \alpha_k$ tales que

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k,$$

por lo tanto,

$$\mathbf{u}\cdot\mathbf{v} = \alpha_1\mathbf{u}\cdot\mathbf{v}_1 + \alpha_2\mathbf{u}\cdot\mathbf{v}_2 + \ldots + \alpha_k\mathbf{u}\cdot\mathbf{v}_k = \mathbf{0}.$$

Es decir, $\mathbf{u}$ es ortogonal a todo $\mathbf{v} \in S$, y por definición, $\mathbf{u}$ es ortogonal a S. □

Ejemplo 49. Verifiquemos que $\mathbf{u} = (-1, -1, 2)^T$ es ortogonal al plano

$$\mathcal{P} = \{(x, y, z)^T,\ x + y - 2z = 0,\ x, y, z \in R\}.$$

Observemos que $\mathcal{P} = \left\{ \begin{pmatrix} 2z - y \\ y \\ z \end{pmatrix},\ y, z \in R \right\} = \left\{ y\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},\ y, z \in R \right\}$; por lo tanto, los vectores $\mathbf{v}_1 = (-1, 1, 0)^T$ y $\mathbf{v}_2 = (2, 0, 1)^T$ generan al plano $\mathcal{P}$. Como el vector $\mathbf{u}$ es ortogonal a los vectores

$\mathbf{v}_1$ y $\mathbf{v}_2$ ($\mathbf{u} \cdot \mathbf{v}_1 = 0$ y $\mathbf{u} \cdot \mathbf{v}_2 = 0$), por el Teorema 28, $\mathbf{u}$ es ortogonal al plano $\mathcal{P}$.[9] □

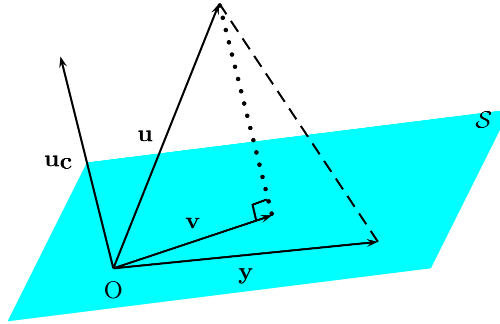
De la misma forma como un vector de R^n se puede descomponer como la suma de dos vectores ortogonales, con uno de ellos en la dirección de otro vector dado, podemos utilizar el concepto de ortogonalidad de un vector a un subespacio para expresar cualquier vector de R^n como la suma de dos vectores ortogonales, con uno de ellos perteneciente al subespacio. Además, este último vector define el punto del subespacio más cercano al punto definido por el vector inicial. Precisemos estas ideas en el siguiente teorema.

Teorema 29 [*Proyección ortogonal sobre un subespacio*].
Sean S un subespacio de R^n y $\mathbf{u}$ un vector arbitrario de R^n. Entonces, existe un único vector $\mathbf{v} \in S$ tal que

1. $\mathbf{u}_c = \mathbf{u} - \mathbf{v}$ es ortogonal al subespacio S ó $\mathbf{u}_c = \mathbf{0}$.

2. $||\mathbf{u}_c|| \leq ||\mathbf{u} - \mathbf{y}||$ para todo $\mathbf{y} \in S$.

Al vector $\mathbf{u}_c$ lo llamamos *la componente de* $\mathbf{u}$ *ortogonal a* S y al vector $\mathbf{v} = Proy_S\mathbf{u}$, la *proyección ortogonal* de $\mathbf{u}$ en S.



Demostración: Sean $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base de S y A la matriz $[\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k]$.

1. Demostrar la existencia de $\mathbf{v} \in S$ tal que $\mathbf{u}_c = \mathbf{u} - \mathbf{v}$ es ortogonal a S es equivalente a demostrar la existencia de un vector $\mathbf{x} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ tal que $\mathbf{u}_c = \mathbf{u} - \mathbf{v} = \mathbf{u} - A\mathbf{x}$ sea ortogonal a S, ya que por definición de base, si $\mathbf{v} \in S$, existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$ tales que $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k$. Ahora, $\mathbf{u}_c = \mathbf{u} - A\mathbf{x}$ es ortogonal a S, si y solo si, $\mathbf{v}_i \cdot (\mathbf{u} - A\mathbf{x}) = 0$, para toda $i = 1, 2, \ldots, n$ o lo que es equivalente, $A^T(\mathbf{u} - A\mathbf{x}) = \mathbf{0}$ o $A^TA\mathbf{x} = A^T\mathbf{u}$. Pero, como A^TA es invertible (ver Ejercicio 11), entonces $A^TA\mathbf{x} = A^T\mathbf{u}$ tiene solución única y por tanto, $\mathbf{v} = A\mathbf{x}$ es el único vector tal que $\mathbf{u}_c = \mathbf{u} - \mathbf{v}$ es ortogonal a S.

2. Sea $\mathbf{y}$ un vector arbitrario de S. Como $Proy_S\mathbf{u} \in S$, entonces $Proy_S\mathbf{u} - \mathbf{y} \in S$. Así,

$$||\mathbf{u} - \mathbf{y}|| = ||\mathbf{u} - Proy_S\mathbf{u} + Proy_S\mathbf{u} - \mathbf{y}|| = ||\mathbf{u}_c + Proy_S\mathbf{u} - \mathbf{y}||$$

Como $\mathbf{u}_c$ es ortogonal a S, en particular, $\mathbf{u}_c$ es ortogonal a $Proy_S\mathbf{u} - \mathbf{y}$. Asi que, por el Teorema de Pitágoras,

$$||\mathbf{u} - \mathbf{y}||^2 = ||\mathbf{u}_c||^2 + ||Proy_S\mathbf{u} - \mathbf{y}||^2.$$

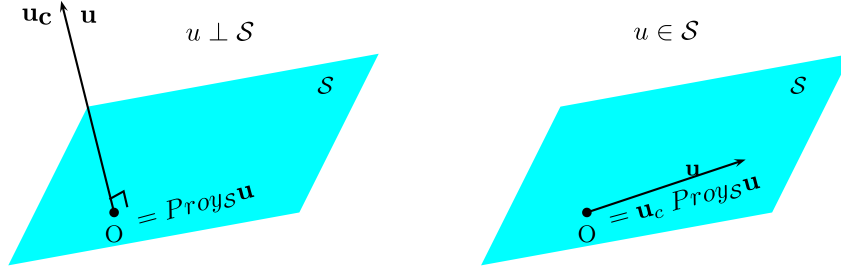
De donde concluimos que

$$||\mathbf{u} - \mathbf{y}||^2 \geq ||\mathbf{u}_c||^2$$

. □

Observemos que si el vector $\mathbf{u}$ es ortogonal a S, entonces $Proy_S\mathbf{u} = \mathbf{0}$ y $\mathbf{u}_c = \mathbf{u}$; y si el vector $\mathbf{u} \in S$, entonces $Proy_S\mathbf{u} = \mathbf{u}$ y $\mathbf{u}_c = \mathbf{0}$.

[9] Usando los conceptos de la última sección del Capítulo 2, también podemos verificar que $\mathbf{u}$ es ortogonal al plano $\mathcal{P}$, observando que $\mathbf{u}$ es paralelo a $\mathbf{n} = (1, 1, -2)^T$, un vector normal del plano $\mathcal{P}$.



Ejemplo 50. Encontremos la proyección ortogonal de $\mathbf{u} = (1, -1, 0, 2)^T$ en el hiperplano

$$\mathcal{H} = \{(x, y, z, w)^T,\ x - z + w = 0,\ x, y, z, w \in R\}.$$

Observemos que $\mathcal{H} = \left\{ \begin{pmatrix} x \\ y \\ x+w \\ w \end{pmatrix},\ x, y, w \in R \right\} = \left\{ x \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + w \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},\ x, y, w \in R \right\}$; por lo tanto, los vectores $\mathbf{v}_1 = (1, 0, 1, 0)^T$, $\mathbf{v}_2 = (0, 1, 0, 0)^T$ y $\mathbf{v}_3 = (0, 0, 1, 1)^T$ generan a $\mathcal{H}$. Como $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto *l.i.* (ver Ejercicio 12), $\mathcal{B}$ es una base de $\mathcal{H}$.

Por el Teorema 29, $Proy_{\mathcal{H}}\mathbf{u} = A\mathbf{x}$, donde $A = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3]$ y $\mathbf{x}$ es la solución de $A^T A\mathbf{x} = A^T\mathbf{u}$. Como

$$A^T A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \qquad A^T\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

y al aplicarle los algoritmos de Eliminación de Gauss y Sustitución hacia atrás a la matriz aumentada $\left(\begin{array}{ccc|c} 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 2 & 2 \end{array}\right)$ obtenemos $\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right)$; por lo tanto, la solución de $A^T A\mathbf{x} = A^T\mathbf{u}$ es $\mathbf{x} = (0, -1, 1)^T$ y

$$Proy_{\mathcal{H}}\mathbf{u} = A\mathbf{x} = (0, -1, 1, 1)^T.$$

□

De los resultados presentados en el teorema anterior, debemos destacar dos aspectos importantes. Primero, el teorema nos garantiza la existencia de un punto en cualquier subespacio de R^n que minimiza la distancia de un punto dado de R^n a los puntos del subespacio y que dicha distancia mínima, la cual llamamos *distancia del punto al subespacio*, es la magnitud de la componente del vector que define el punto dado, ortogonal al subespacio, lo cual es muy importante en los procesos de aproximación por Mínimos Cuadrados Lineales[10].

Segundo, observemos que, aunque no es necesario, si en la demostración anterior, hubiésemos escogido una base ortogonal del subespacio, por el Teorema 26, $A^T A = D$; o si hubiésemos escogido una base ortonormal, por el mismo Teorema 26, $A^T A = I$, el cálculo de $Proy_S\mathbf{u}$ se hubiese simplificado. En efecto, si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base ortogonal de S,

$$Proy_S\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{u} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 + \ldots + \frac{\mathbf{u} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k}\mathbf{v}_k, \tag{4.11}$$

o si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base ortonormal de S,

$$Proy_S\mathbf{u} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2 + \ldots + (\mathbf{u} \cdot \mathbf{v}_k)\mathbf{v}_k, \tag{4.12}$$

[10] El problema de Mínimos Cuadrados Lineales consiste en hallar el vector $\mathbf{x} \in R^n$ tal que minimice $||A\mathbf{x} - \mathbf{b}||$, para una matriz A de tamaño $m \times n$ y un vector $\mathbf{b}$ dados. En este mismo contexto, el sistema de ecuaciones lineales, $A^T A\mathbf{x} = A^T\mathbf{u}$, el cual proporciona la(s) solución(es) del problema de mínimos cuadrados lineales, se conoce con el nombre de *Ecuaciones Normales*.

(ver Ejercicio 24). Sin embargo, no lo hicimos así para no dar la falsa impresión que la proyección ortogonal requiere de una base ortogonal y porque aún no hemos demostrado que todo subespacio de R^n tiene una base ortogonal.

Ejemplo 51. Calculemos la distancia del punto definido por el vector $\mathbf{u}$ al hiperplano $\mathcal{H}$, para $\mathbf{u}$ y $\mathcal{H}$ dados en el Ejemplo 50.

Por la observación anterior, la distancia de $\mathbf{u}$ a $\mathcal{H}$ es la norma de $\mathbf{u}_c = \mathbf{u} - Proy_{\mathcal{H}}\mathbf{u}$, la componente de $\mathbf{u}$ ortogonal a $\mathcal{H}$; es decir,

$$||\mathbf{u}_c|| = ||\mathbf{u} - Proy_{\mathcal{H}}\mathbf{u}|| = ||(1, 0, -1, 1)^T|| = \sqrt{3}.$$

□

Ejemplo 52. Encontremos la proyección ortogonal de $\mathbf{w}$ en V, para el vector $\mathbf{w}$ y el subespacio V dados en el Ejemplo 47.

Teniendo en cuenta los datos y la notación del Ejemplo 47, como $\mathcal{B}$ es una base ortonormal de V, por la Ecuación (4.12), tenemos que

$$\begin{aligned} Proy_V\mathbf{w} &= (\mathbf{w}\cdot\mathbf{v}_1)\,\mathbf{v}_1 + (\mathbf{w}\cdot\mathbf{v}_2)\,\mathbf{v}_2 \\ &= \left[\frac{-3}{\sqrt{3}}\right]\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\-1\end{pmatrix} + \left[\frac{4}{\sqrt{14}}\right]\frac{1}{\sqrt{14}}\begin{pmatrix}2\\1\\3\end{pmatrix} \\ &= \begin{pmatrix}-1\\-1\\1\end{pmatrix} + \frac{1}{7}\begin{pmatrix}4\\2\\6\end{pmatrix} = \frac{1}{7}\begin{pmatrix}-3\\-5\\13\end{pmatrix}. \end{aligned}$$

□

Observemos también que como $Proy_S\mathbf{u} \in S$ y $\mathbf{u}_c = \mathbf{u} - Proy_S\mathbf{u}$, si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base de S, entonces

$$Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{u}\} = Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{u}_c\}. \tag{4.13}$$

Con base en los resultados anteriores, veamos como obtener una base ortogonal y una base ortonormal para cualquier subespacio de R^n, a partir de una base del subespacio, y demostrar así la existencia de bases ortogonales y ortonormales para todo subespacio de R^n.

Teorema 30 [*Proceso de ortogonalización de Gram-Schmidt*].
Todo subespacio S de R^n, $S \neq \{\mathbf{0}\}$, tiene al menos una base ortonormal.

Demostración: Sea S un subespacio de R^n diferente de $\{\mathbf{0}\}$. Por el Teorema 6, existe al menos una base de S. Sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base de S. Construyamos, a partir de esta base, a $\mathcal{B}' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$, una base ortogonal de S y luego la normalizamos para obtener una base ortonormal.

La idea de la demostración es definir a $\mathbf{u}_i$ como la componente de $\mathbf{v}_i$ ortogonal al subespacio generado por los vectores $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}$, previamente definidos, empezando con $\mathbf{u}_1 = \mathbf{v}_1$ y el subespacio $Gen\{\mathbf{u}_1\}$. Así, $\mathbf{u}_2$ es la componente de $\mathbf{v}_2$ ortogonal al subespacio $Gen\{\mathbf{u}_1\}$, por lo tanto,

$$\mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1$$

y por (4.13),

$$Gen\{\mathbf{v}_1, \mathbf{v}_2\} = Gen\{\mathbf{u}_1, \mathbf{v}_2\} = Gen\{\mathbf{u}_1, \mathbf{u}_2\}.$$

Supongamos que de esta forma, hemos construido el conjunto de vectores ortogonal $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}\}$ tal que $Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_{i-1}\} = Gen\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}\}$. Definamos $\mathbf{u}_i$ como la componente de $\mathbf{v}_i$ ortogonal al subespacio $Gen\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}\}$. Por lo tanto,

$$\mathbf{u}_i = \mathbf{v}_i - \left(\frac{\mathbf{v}_i\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{v}_i\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \ldots + \frac{\mathbf{v}_i\cdot\mathbf{u}_{i-1}}{\mathbf{u}_{i-1}\cdot\mathbf{u}_{i-1}}\mathbf{u}_{i-1}\right)$$

y por (4.13),

$$\begin{aligned} Gen\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_{i-1}, \mathbf{v}_i\} &= Gen\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{v}_i\} \\ &= Gen\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_i\}, \end{aligned}$$

lo cual podemos continuar hasta $i = k$, obteniendo asi un conjunto ortogonal de k vectores que genera a S y por consiguiente, es una base ortogonal de S. Finalmente, es fácil verificar que

$$\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \frac{\mathbf{u}_2}{||\mathbf{u}_2||}, \ldots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$$

es una base ortonormal de S. □

Ejemplo 53. Encontremos una base ortogonal y otra ortonormal del hiperplano $\mathcal{H}$ dado en el Ejemplo 50.

En el Ejemplo 50, demostramos que los vectores $\mathbf{v}_1 = (1,0,1,0)^T$, $\mathbf{v}_2 = (0,1,0,0)^T$ y $\mathbf{v}_3 = (0,0,1,1)^T$ forman una base de $\mathcal{H}$. Siguiendo las ideas sugeridas en la demostración del Teorema 30 para construir una base ortogonal, sean $\mathbf{u}_1 = \mathbf{v}_1$,

$$\mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{0}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

y

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{0}{1}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

Finalmente, si normalizamos los vectores de $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, la cual es una base ortogonal de $\mathcal{H}$, obtenemos

$$\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix} \right\},$$

que es una base ortonormal de $\mathcal{H}$. □

4.10. Factorización QR

De la misma forma como el algoritmo de Eliminación de Gauss nos permitió factorizar una matriz como el producto de dos matrices triangulares, una inferior y otra superior, el Proceso de ortogonalización de Gram-Schmidt (Teorema 30) nos permite factorizar una matriz como el producto de dos matrices, una ortonormal y otra triangular superior. Aunque el resultado es válido para cualquier matriz de tamaño $m \times n$, en el siguiente teorema, nos limitamos al caso de las matrices de *rango completo* (matrices cuyo rango es igual al número de columnas; es decir, matrices cuyas columnas forman un conjunto *l.i.*)

Teorema 31 [*Factorización QR*].
Para toda matriz A de tamaño $m \times n$, cuyas columnas forman un conjunto *l.i.*, existe una matriz Q de tamaño $m \times n$, cuyas columnas forman un conjunto ortonormal, y una matriz triangular superior R de tamaño $n \times n$ tales que

$$A = QR.$$

Demostración: Sea $A = [\mathbf{a}_1\ \mathbf{a}_2\ \dots\ \mathbf{a}_n]$, cuyas columnas $\mathbf{a}_j \in R^m,\ j = 1, 2, \dots, n$, forman un conjunto *l.i.* Por el Teorema 18, $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ es una base de C_A, el espacio columna de A. Por el Teorema 30, existe $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$, una base ortogonal de C_A, tal que $\mathbf{u}_1 = \mathbf{a}_1$ y

$$\mathbf{u}_j = \mathbf{a}_j - \left(\frac{\mathbf{a}_j \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{a}_j \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \dots + \frac{\mathbf{a}_j \cdot \mathbf{u}_{j-1}}{\mathbf{u}_{j-1} \cdot \mathbf{u}_{j-1}}\mathbf{u}_{j-1} \right),$$

para $j = 2, \dots, n$. Reescribiendo estas ecuaciones, tenemos que $\mathbf{a}_1 = \mathbf{u}_1$ y

$$\mathbf{a}_j = \left(\frac{\mathbf{a}_j \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{a}_j \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \dots + \frac{\mathbf{a}_j \cdot \mathbf{u}_{j-1}}{\mathbf{u}_{j-1} \cdot \mathbf{u}_{j-1}}\mathbf{u}_{j-1} \right) + \mathbf{u}_j,$$

para $j = 2, \dots, n$. Si $\hat{Q} = [\mathbf{u}_1\ \mathbf{u}_2\ \dots\ \mathbf{u}_n]$, las anteriores ecuaciones las podemos escribir como

$$\mathbf{a}_j = \hat{Q}\mathbf{r}_j, \quad \text{para } j = 1, 2, \dots n,$$

donde r_{ij}, la componente i de $\mathbf{r}_j$, es

$$r_{ij} = \begin{cases} \frac{\mathbf{a}_j \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} & \text{si } i < j \\ 1 & \text{si } i = j \\ 0 & \text{si } i > j. \end{cases} \quad \text{para } i = 1, 2, \dots, n. \tag{4.14}$$

Si $\hat{R} = [\mathbf{r}_1\ \mathbf{r}_2\ \dots\ \mathbf{r}_n]$, tenemos que $A = \hat{Q}\hat{R}$. Sea D la matriz diagonal con $d_{ii} = ||\mathbf{u}_i||$, la cual es invertible, ya que $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ es un conjunto ortogonal y sean $Q = \hat{Q}D^{-1}$ y $R = D\hat{R}$. Las matrices que buscamos son Q y R, ya que

1. La matriz Q tiene columnas ortonormales: Por el Teorema 26, $\hat{Q}^T\hat{Q} = D^2$; por lo tanto,
$$Q^TQ = (\hat{Q}D^{-1})^T(D^{-1}\hat{Q}) = D^{-1}(\hat{Q}^T\hat{Q})D^{-1} = D^{-1}D^2D^{-1} = I_n$$
2. La matriz R es triangular superior: Porque $\hat{R}$ es triangular superior y D es diagonal.
3. $A = QR$: Por la definición de Q y R, tenemos que
$$QR = (\hat{Q}D^{-1})(D\hat{R}) = \hat{Q}(D^{-1}D)\hat{R} = \hat{Q}\hat{R} = A$$

□

Ejemplo 54. Calculemos la factorización QR de la matriz $A = \begin{pmatrix} 2 & 2 & 0 \\ -2 & 0 & 6 \\ 0 & 2 & 0 \\ 1 & 5 & 3 \end{pmatrix}$.

Utilizando el Proceso de Ortogonalización de Gram-Schmidt para hallar $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, una base ortogonal de C_A, a partir del conjunto de vectores formados por las columnas de A, obtenemos $\mathbf{u}_1 = (2, -2, 0, 1)^T$, $\mathbf{u}_2 = (0, 2, 2, 4)^T$ y $\mathbf{u}_3 = (2, 2, -2, 0)^T$. Por el resultado del Teorema 31, si $\hat{Q} = [\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3]$, $\hat{R} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ y $D = \begin{pmatrix} ||u_1|| & 0 & 0 \\ 0 & ||u_2|| & \\ 0 & 0 & ||u_3|| \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2\sqrt{6} & 0 \\ 0 & 0 & 2\sqrt{3} \end{pmatrix}$, entonces $A = QR$, donde

$$Q = \hat{Q}D^{-1} = \left[\frac{\mathbf{u}_1}{||\mathbf{u}_1||}\ \frac{\mathbf{u}_2}{||\mathbf{u}_2||}\ \frac{\mathbf{u}_3}{||\mathbf{u}_3||} \right] = \begin{pmatrix} 2/3 & 0 & 1/\sqrt{3} \\ -2/3 & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 1/\sqrt{6} & -1/\sqrt{3} \\ 1/3 & 2/\sqrt{6} & 0 \end{pmatrix} \text{ y } R = D\hat{R} = \begin{pmatrix} 3 & 3 & -3 \\ 0 & 2\sqrt{6} & 2\sqrt{6} \\ 0 & 0 & 2\sqrt{3} \end{pmatrix}.$$

□

Un análisis detallado del Teorema 31 nos lleva a concluir que una vez calculada una base ortogonal a partir de una base dada, tenemos todos los elementos necesarios para calcular la factorización QR de la matriz cuyas columnas son los vectores de la base dada. En efecto, la matriz Q tendrá como columnas los vectores de la base ortogonal hallada debidamente normalizados (en otras palabras, los vectores de la base ortonormal correspondiente) y la matriz R será el producto de la matriz diagonal formada con las normas de los vectores de la base ortogonal, cuyos valores se calcularon para normalizar la base ortogonal, y la matriz triangular superior unitaria cuyas columnas están formadas por el opuesto de los coeficientes utilizados en el cálculo de cada uno de los vectores de la base ortogonal, como lo define la expresión (4.14). Veamos un ejemplo.

Ejemplo 55. Utilicemos los resultados del Ejemplo 53 para calcular la factorización QR de la matriz cuyas columnas son los vectores $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$; es decir, de la matriz $A = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3]$.

Teniendo en cuenta el análisis anterior, sea Q la matriz cuyas columnas son los vectores de la base ortonormal calculada (es decir, $Q = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/2\sqrt{6} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/2\sqrt{6} \\ 0 & 0 & 1/\sqrt{6} \end{pmatrix}$); D la matriz diagonal formada con las normas de los vectores de la base ortogonal (es decir, $D = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{6} \end{pmatrix}$) y $\hat{R}$ la matriz triangular superior unitaria cuyas columnas son los opuestos de los coeficientes utilizados en el cálculo de los vectores de la base ortogonal (es decir, $\hat{R} = \begin{pmatrix} 1 & 0 & 0{,}5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$). Finalmente, $R = D\hat{R} = \begin{pmatrix} \sqrt{2} & 0 & \sqrt{2}/2 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{6} \end{pmatrix}$ y podemos verificar que

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/2\sqrt{6} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/2\sqrt{6} \\ 0 & 0 & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & \sqrt{2}/2 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{6} \end{pmatrix} = QR.$$

□

Para terminar, observemos que la factorización QR, además de darnos una base ortonormal de C_A, el espacio columna de la matriz A, nos brinda una forma de resolver $A^T A\mathbf{x} = A^T\mathbf{b}$, el sistema de ecuaciones lineales que surge en el cálculo de la solución del Problema de Mínimos Cuadrados Lineales, ($Min\ ||A\mathbf{x} - \mathbf{b}||$), sin necesidad de calcular la matriz $A^T A$, lo cual se deja como ejercicio para el lector.

La factorización QR, al igual que la factorización LU, también nos permite resolver sistemas de ecuaciones lineales, en general. En efecto, para resolver el sistema $A\mathbf{x} = \mathbf{b}$, como $A = QR$ y $Q^T Q = I$, tenemos $(QR)\mathbf{x} = \mathbf{b}$, o equivalentemente, $Q(R\mathbf{x}) = \mathbf{b}$; por lo tanto, $R\mathbf{x} = Q^T\mathbf{b}$, el cual es un sistema de ecuaciones lineales que se puede resolver mediante sustitución hacia atrás, ya que la matriz R es triangular superior[11].

4.11. Ejercicios

1. Determine si los siguientes conjuntos, con las operaciones indicadas, son espacios vectoriales (reales).

   a) El conjunto de los números complejos con la suma $(a + bi) + (c + di) = (a + c) + (b + d)i$ y el producto por escalar $\lambda(a + bi) = (\lambda a) + (\lambda b)i$.

[11] Ver más detalles en MARTINEZ H.J y PEREZ R.,Introducción al Álgebra Lineal Numérica. Ed. Universidad del Cauca, 1990.

b) El conjunto de vectores de R^2 con la suma $\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 \\ b_1 + b_2 \end{pmatrix}$ y el producto por escalar $\lambda \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ b_1 \end{pmatrix}$.

2. Verifique los Axiomas 2, 3, 7, 8 y 10 en el Ejemplo 1.

3. Determine si los siguientes conjuntos son espacios vectoriales (reales). (AYUDA: Determine si los conjuntos dados son subespacios de un espacio vectorial conocido)

   a) El conjunto de matrices simétricas 8×8.

   b) El conjunto de puntos del segundo cuadrante del plano cartesiano.

   c) El conjunto de polinomios de grado 3.

   d) El conjunto de puntos de la recta que pasa por $P = (1, -2, 3)^T$ y $Q = (5, 0, -1)^T$.

   e) El conjunto de puntos del plano que pasa por $P = (1, -2, 3, -1)^T$, $Q = (5, 0, -1, 2)^T$ y $R = (3, 4, -7, 0)^T$.

   f) El conjunto de números racionales.

   g) El conjunto de matrices triangulares superiores 20×20

   h) El conjunto de funciones de R en R, derivables.

   i) El conjunto de polinomios de grado menor o igual a 4, tales que evaluados en cero dan 1.

4. Verifique los Axiomas 2, 3, 7, 8, 9 y 10 en el Ejemplo 3.

5. Geométricamente, cuales son los subespacios de R^2? de R^3? de R^4? (En R^n, para $n \geq 5$, existen otros subespacios distintos a los de R^4?)

6. En cada caso, determine si el vector $\mathbf{v}$ es combinación lineal del conjunto de vectores S y en caso de serlo, encuentre los coeficientes de la combinación lineal; diga además si la combinación es única. Que puede decir del conjunto de vectores dado?

   a) $S = \left\{ \begin{pmatrix} 2 \\ 4 \\ -5 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ -2 \\ -5 \end{pmatrix} \right\}; \quad \mathbf{v} = \begin{pmatrix} -5 \\ 5 \\ 0 \end{pmatrix}$

   b) $S = \left\{ \begin{pmatrix} 3 \\ 4 \\ -7 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ -4 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix} \right\}; \quad \mathbf{v} = \begin{pmatrix} 0 \\ -2 \\ 2 \\ -3 \end{pmatrix}$

   c) $S = \{1 - x + x^2,\ 2 + x^2,\ -1 + 2x\}, \quad \mathbf{v} = 3 - 2x + 2x^2$.

   d) $S = \{1 - x + x^2 - x^3,\ 2 + x^2,\ 3 + x + x^2 + x^3\}, \quad \mathbf{v} = 3 - 2x + 2x^2 - x^3$.

   e) $S = \left\{ \begin{pmatrix} 3 & 4 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -3 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \right\}; \quad \mathbf{v} = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$

   f) $S = \left\{ \begin{pmatrix} 3 & 4 \\ -7 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 2 \\ -4 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix} \right\}; \quad \mathbf{v} = \begin{pmatrix} 0 & -2 \\ 2 & -3 \end{pmatrix}$

7. Para cada uno de los conjuntos S del ejercicio anterior, encuentre el conjunto generado por S. Existe un conjunto de vectores con menor número de ellos que genere el mismo conjunto?

8. Verifique que para todo trio de vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ de un espacio vectorial V,

   a) Gen$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$=Gen$\{\mathbf{u}, \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{w}\}$.

   b) Si $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ es *l.i.*, entonces $\{\mathbf{u} - \mathbf{v},\ \mathbf{v} - \mathbf{w},\ \mathbf{u} + \mathbf{w}\}$ también es *l.i.*

9. Encuentre un conjunto generador, un conjunto *l.d.* y un conjunto *l.i.* de cada uno de los siguientes espacios vectoriales.

   a) $C = \{z = a + bi : \ a, b \in R, \ i = \sqrt{-1}\}$.

   b) El hiperplano $\mathcal{H} : \ 3x - 2y + w = 0$ de R^4.

   c) $N_A = \{\mathbf{x} \in R^3 : \ A\mathbf{x} = 0\}, \quad A = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$

   d) $S = Gen \left\{ \begin{pmatrix} 5 \\ 0 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} \begin{pmatrix} 9 \\ -3 \\ -2 \\ 4 \\ 6 \end{pmatrix} \right\}$

   e) $S = Gen\left\{3x - 2x^2, \ 2 + x, \ -4 + x - 2x^2\right\}$

   f) $S = \{A \in \mathcal{M}_{3\times 3} : \ A = A^T\}$

   g) $S = \{A = (a_{ij})_{3\times 3} : \ a_{ij} = 0, i \neq j\}$

10. Determine si cada uno de los siguientes conjuntos es *l.i.* En caso negativo, encuentre un vector que sea combinación lineal de los restantes.

   a) $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

   b) $\left\{3x - 2x^2, \ 2 + x, \ 2x + 2x^2\right\}$.

   c) $\left\{ \begin{pmatrix} 1 & 0 \\ -3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -3 & -3 \end{pmatrix} \right\}$.

   d) $\left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} \right\}$.

11. Demuestre que si $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base del espacio vectorial V y A es la matriz $[\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k]$, entonces las columnas de AA^T forman un conjunto *l.i.* y por tanto la matriz AA^T es invertible.

12. Demuestre que el conjunto $\mathcal{B}$ del Ejemplo 50 es *l.i.*

13. Encuentre el subespacio más pequeño que contiene cada uno de los conjuntos de vectores dados.

   a) $\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 2 \\ 3 \end{pmatrix} \right\}$

   b) $\left\{3x - 2x^2, \ 2 + x, \ 2x + 2x^2, x^3, \ -2x - 2x^2 + 2x^3\right\}$

   c) $\left\{ \begin{pmatrix} 1 & -3 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} -3 & 2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -2 & 2 \\ 2 & 0 \end{pmatrix} \right\}$

14. Determine si cada una de las siguientes afirmaciones es FALSA o VERDADERA y diga por qué

   a) Si el determinante de una matriz 5×5 es 3, la matriz tiene máximo 3 columnas que forman un conjunto *l.i.*

   b) Si $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ es un conjunto de vectores *l.i.*, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ también es un conjunto de vectores *l.i.*

   c) Si $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ es un conjunto de vectores *l.d.*, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ también es un conjunto de vectores *l.d.*

d) Si $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ es un conjunto de vectores *l.i.*, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$ también es un conjunto de vectores *l.i.*

e) Si $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ es un conjunto de vectores *l.d.*, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$ también es un conjunto de vectores *l.d.*

f) Un subconjunto finito diferente de $\{0\}$ no puede ser un subespacio vectorial.

g) La unión de dos subespacios vectoriales de un mismo espacio vectorial es un subespacio vectorial.

h) La intersección de dos subespacios vectoriales es un subespacio vectorial.

i) Si $Gen\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\} = Gen\{\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_4\}$, $\{\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_4\}$ es *l.i.*

j) Si $Gen\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\} = Gen\{\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_4\}$, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ es *l.d.*

k) El conjunto de matrices antisimétricas 3×3 es un subespacio de $\mathcal{M}_{3\times 3}$

l) El conjunto de puntos dentro de un circulo alrededor del origen de radio 1 es un subespacio de R^2.

m) El conjunto de matrices triangulares inferiores con unos en la diagonal 5×5 es un subespacio de $\mathcal{M}_{5\times 5}$.

n) El conjunto de matrices escalares 4×4 es un subespacio de $\mathcal{M}_{4\times 4}$.

ñ) El conjunto de matrices elementales 10×10 es un subespacio de $\mathcal{M}_{10\times 10}$.

o) El conjunto de polinomios de grado menor o igual a 5, con coeficientes enteros, es un subespacio de $\mathcal{P}_5$.

p) El conjunto de polinomios de grado igual a 3 es un subespacio de $\mathcal{P}_3$.

15. En cada caso, determine si el conjunto $\mathcal{B}$ forma una base del espacio vectorial V.

a) $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix} \right\}$, $V = R^3$.

b) $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} \right\}$, $V = R^3$.

c) $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} \right\}$, $V = R^3$.

d) $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix}, \begin{pmatrix} -5 \\ 0 \\ 5 \end{pmatrix} \right\}$, $V = R^3$.

e) $\mathcal{B} = \{1 + x, 1 - x, x^2\}$, $V = \mathcal{P}_2$.

f) $\mathcal{B} = \{1 + x, x^2, 1 - x^3\}$, $V = \mathcal{P}_3$.

g) $\mathcal{B} = \left\{ \begin{pmatrix} 1 & -3 \\ 5 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 5 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \right\}$, $V = \mathcal{M}_{2\times 2}$.

h) $\mathcal{B} = \{1 + x,\ x^2\}$, $V = \{a_0 + a_1x + a_2x^2 \in \mathcal{P}_2 : a_1 = 0\}$.

i) $\mathcal{B} = \{1,\ 1 + x^2\}$, $V = \{a_0 + a_1x + a_2x^2 \in \mathcal{P}_2 : a_1 = 0\}$.

16. Encuentre una base del espacio vectorial dado y determine su dimensión.

a) $H = \{(x\ y\ z\ w)^T : x + y + z + w = 0\}$.

b) $W = \left\{ \begin{pmatrix} a & b \\ -b & -a \end{pmatrix} : a, b \in R \right\}$.

c) $K = \{a_0 + a_1x + a_2x^2 + a_3x^3 \in \mathcal{P}_3 : a_0 - 2a_1 = 0\}$.

d) El conjunto de matrices simétricas 3×3.

e) $Gen\{1 + x, x^2, 1 - x^3\}$.

f) $Gen\{1 + x, 1 - x, 1 + x^2, 1 - x^2\}$.

17. A partir del conjunto S dado, construya una base del espacio vectorial H que contenga o esté contenida en S.

a) $S = \{(1\ 2\ 3\ 4)^T, (1\ 3\ 5\ 7)^T, (1\ 2\ 4\ 6)^T\}$, $H = R^4$.

b) $S = \left\{ \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \right\}$, $H = \mathcal{M}_{2\times 2}$.

c) $S = \{1 - x^2, 1 + x\}$, $H = \mathcal{P}_2$.

d) $S = \{1 - x^2, 1 + x, 1 - x^3, 1 + x^3, x^2 + x^3\}$, $H = \mathcal{P}_3$.

e) $S = \{(2\ 3 - 2 - 3)^T, (1\ 2\ 3 - 6)^T\}$, $H = \{(x\ y\ z\ w)^T : x + y + z + w = 0\}$.

18. Justifique que $\mathcal{B}$ y $\mathcal{B}'$ son bases de V, calcule el vector $\mathbf{u}$ cuyas coordenadas en una de las dos bases se dan y calcule las coordenadas del vector $\mathbf{u}$ en la otra base.

a) $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} \right\}$, $\mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$, $V = R^3$,
$[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$.

b) $\mathcal{B} = \left\{ \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$,
$\mathcal{B}' = \left\{ \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$, $V = \mathcal{M}_{2\times 2}$, $[\mathbf{u}]_{\mathcal{B}'} = \begin{pmatrix} 0 \\ 1 \\ -2 \\ -1 \end{pmatrix}$

c) $\mathcal{B} = \{2 - x, 1 - x^2, 1 + x\}$, $\mathcal{B}' = \{1 + x, 2 - x, 1 - x^2\}$, $V = \mathcal{P}_2$, $[\mathbf{u}]_{\mathcal{B}'} = \begin{pmatrix} -3 \\ 0 \\ -2 \end{pmatrix}$

d) $\mathcal{B} = \{1, x, x^2, x^3\}$, $\mathcal{B}' = \{1 + x, x + x^2, x^2 + x^3, 1 - x^3\}$, $V = \mathcal{P}_3$, $[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \\ -2 \\ -1 \end{pmatrix}$

e) $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix} \right\}$, $\mathcal{B}' = \left\{ \begin{pmatrix} -3 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} \right\}$, $V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x + y + z = 0 \right\}$,
$[\mathbf{u}]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

19. Calcule el rango y la nulidad para cada una de las matrices dadas.

$$A = \begin{pmatrix} -1 & 3 & -5 \\ 0 & 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & -1 & 3 \\ 0 & 1 & 2 \\ -4 & 2 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & -1 & 3 \\ 0 & 1 & 2 \\ -4 & 2 & 1 \\ 2 & 2 & -1 \end{pmatrix}$$

20. Demuestre que el conjunto de vectores formado por las filas, diferentes de $\mathbf{0}$, de una matriz escalonada es $l.i.$

21. Verifique si los siguientes conjuntos son ortogonales y si son ortonormales.

a) $S = \left\{ \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$ b) $S = \left\{ \begin{pmatrix} \frac{1}{2} \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix} \right\}.$

c) $S = \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}.$ d) $S = \left\{ \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\}.$

22. Calcule la proyección de $\mathbf{u}$ en el subespacio H y la componente de $\mathbf{u}$ ortogonal a H.

a) $\mathbf{u} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, H = Gen \left\{ \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right\}.$

b) $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, H = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x + y + z = 0, \right\}.$

c) $\mathbf{u} = (1\ 0\ 2 - 1),\ H = \left\{ (x\ y\ z\ w)^T : 2x - y + w = 0, \right\}.$

23. Construya una base ortogonal y una base ortonormal de los espacios vectoriales H del ejercicio anterior.

24. Demuestre que si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base ortogonal de S, entonces

$$Proy_S\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{u}\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 + \ldots + \frac{\mathbf{u}\cdot\mathbf{v}_k}{\mathbf{v}_k\cdot\mathbf{v}_k}\mathbf{v}_k,$$

y si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base ortonormal de S, entonces

$$Proy_S\mathbf{u} = (\mathbf{u}\cdot\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}\cdot\mathbf{v}_2)\mathbf{v}_2 + \ldots + (\mathbf{u}\cdot\mathbf{v}_k)\mathbf{v}_k.$$

25. Determine si cada una de las siguientes afirmaciones es FALSA o VERDADERA y diga por qué

a) La dimensión del espacio de las matrices diagonales 4×4 es 4.

b) La dimensión del espacio de los polinomios de grado menor o igual a 4, que evaluados en 1 es 0, es 3.

c) La dimensión de un hiperplano en R^5 es 4.

d) La dimensión de $Gen\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ es 4.

e) La dimensión de $Gen\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$, cuando $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ es $l.i.$, es 4.

f) Las coordenadas de una matriz 3×5 en una base de $\mathcal{M}_{3\times5}$ es un vector de R^8.

g) Las coordenadas de un vector de un plano (en R^5) en una base del plano es un vector de R^2.

h) Las coordenadas de un vector de un hiperplano (en R^5) en una base de R^5 es un vector de R^4.

i) Si $[\mathbf{u}]_{\mathcal{B}} \in R^5$, entonces $dim(Gen\ \mathcal{B}) = 5$.

j) Dadas P y Q, las matrices de transición de $\mathcal{B}$ a $\mathcal{B}''$ y de $\mathcal{B}'$ a $\mathcal{B}''$, respectivamente, la ecuación $P[\mathbf{u}]_{\mathcal{B}} = Q[\mathbf{u}]_{\mathcal{B}'}$ permite calcular las coordenadas del vector $\mathbf{u}$ en una base, conociendo las coordenadas del vector $\mathbf{u}$ en la otra base.

k) Un conjunto de 5 matrices 3×2 puede generar a $\mathcal{M}_{3\times2}$

l) Un conjunto de 5 polinomios de grado menor o igual 3 genera a $\mathcal{P}_3$

m) Un conjunto de 5 vectores de un hiperplano en R^5 que pasa por el origen puede ser *l.i.*

n) Un conjunto de 5 matrices diagonales 6×6 es *l.i.*

26. Determine si cada una de las siguientes afirmaciones es FALSA o VERDADERA y diga por qué

a) Si $S = \{\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{t}\} \subset \mathcal{P}_2$, S puede ser un conjunto ortogonal.

b) Si $S = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\} \subset V$ es un conjunto ortonormal, entonces $dimV \geq 4$.

c) Es posible encontrar un conjunto ortogonal de 3 vectores de un hiperplano en R^6.

d) Para calcular la proyección ortogonal de un vector en un subespacio, se requiere una base ortogonal del subespacio.

e) La proyección ortogonal de un vector sobre un subespacio es ortogonal al subespacio.

f) La suma de la proyección ortogonal de un vector en un subespacio con la componente del vector ortogonal al subespacio es el vector.

g) Si $\nu(A) = 0$ y A es una matriz 7×4, el sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$ tiene solución única para todo vector $\mathbf{b}$.

h) Una matriz A de tamaño 4×7 no puede tener una nulidad igual a cero.

i) Si el rango de una matriz 5×8 es 5, la nulidad de su transpuesta es 3.

j) El rango de una matriz 5×8 no puede ser 6.

k) Una base del espacio columna de una matriz es la conformada por las columnas pivotales de una matriz escalonada equivalente.

l) Una base del espacio fila de una matriz es la conformada por las filas que tienen pivotes en una matriz escalonada equivalente.

m) La dimensión del espacio fila de una matriz es igual al rango de la matriz.

n) Si $\rho(A) = 5$ y A es una matriz 5×9, el sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$ tiene infinitas soluciones para todo vector $\mathbf{b}$.

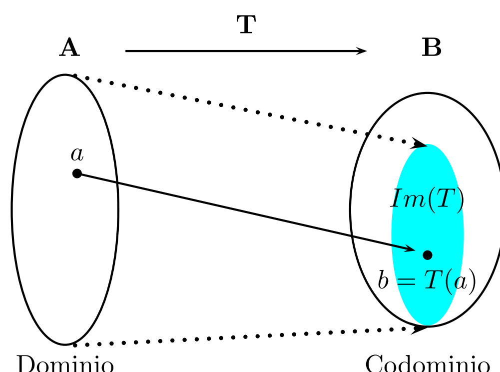
27. Muestre cómo resolver $A^T A\mathbf{x} = A^T\mathbf{b}$ sin calcular $A^T A$, utilizando la factorización QR de A.

Capítulo 5

TRANSFORMACIONES LINEALES

5.1. Introducción

Recordemos que una *función* $T : A \longrightarrow B$ es una "regla de asociación" entre los elementos de A y los elementos de B, tal que a cada elemento a de A se le asocia un único elemento b de B al que le llamamos *imagen de* a por medio de T y denotamos $b = T(a)$. A los conjuntos A y B les llamamos *dominio* y *codominio* de T, respectivamente, y al subconjunto de B formado por todas las imágenes de los elementos de A lo llamamos *conjunto imagen* de T y lo denotamos $Im(T)$.



En este capítulo, estamos interesados en el estudio de las funciones entre espacios vectoriales que sean "compatibles" con las operaciones de suma y producto por escalar definidas en cada uno de ellos; es decir, que la imagen de una suma de vectores sea la suma de las imágenes y que la imagen de un producto por escalar de un vector sea también un producto por escalar de la imagen del vector.

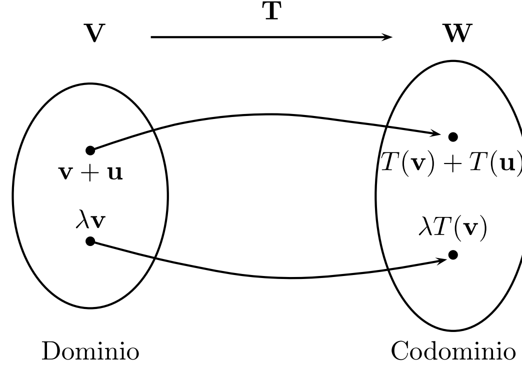
5.2. Definición y Propiedades Básicas

Precisemos la idea planteada en la introducción con la siguiente definición.

Definición 1 [*Transformación lineal*]. Dados dos espacios vectoriales V y W, diremos que la función $T : V \longrightarrow W$ es una *transformación lineal* de V en W, si y solo si,

1. $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ para todo $\mathbf{v}_1, \mathbf{v}_2 \in V$. *(Propiedad aditiva)*

2. $T(\lambda \mathbf{v}_1) = \lambda T(\mathbf{v}_1)$ para todo $\mathbf{v}_1 \in V$ y todo $\lambda \in R$. *(Propiedad homogénea)*



Observación: Es importante aclarar que estamos denotando de la misma forma la suma y el producto por escalar definidos tanto en el espacio vectorial V como en el espacio vectorial W, así sean operaciones diferentes. Igualmente, al vector $\mathbf{0}$ de V y al de W los denotamos igual, así sean vectores diferentes.

Ejemplo 1. Consideremos la función $T : R^3 \longrightarrow R^2$, tal que $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x-2y \\ 2z-x \end{pmatrix}$. Verifiquemos que T es una transformación lineal.

Veamos que T satisface la *propiedad aditiva*. Sean $(x_1, y_1, z_1)^T$ y $(x_2, y_2, z_2)^T$ vectores de R^3, entonces

$$\begin{aligned} T\left[\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right] &= T\begin{pmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} \\ &= \begin{pmatrix} (x_1+x_2)-2(y_1+y_2) \\ 2(z_1+z_2)-(x_1+x_2) \end{pmatrix} \end{aligned}$$

$$\begin{aligned} T\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} &= \begin{pmatrix} x_1-2y_1 \\ 2z_1-x_1 \end{pmatrix} + \begin{pmatrix} x_2-2y_2 \\ 2z_2-x_2 \end{pmatrix} \\ &= \begin{pmatrix} (x_1-2y_1)+(x_2-2y_2) \\ (2z_1-x_1)+(2z_2-x_2) \end{pmatrix} \\ &= \begin{pmatrix} (x_1+x_2)-2(y_1+y_2) \\ 2(z_1+z_2)-(x_1+x_2) \end{pmatrix}; \end{aligned}$$

de donde,

$$T\left[\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right] = T\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}.$$

Para verificar que T satisface la *propiedad homogénea*, tomemos el escalar λ y el vector $(x, y, z)^T$. Entonces,

$$\begin{aligned} T\left[\lambda\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] &= T\begin{pmatrix} \lambda x \\ \lambda y \\ \lambda z \end{pmatrix} \\ &= \begin{pmatrix} (\lambda x)-2(\lambda y) \\ 2(\lambda z)-(\lambda x) \end{pmatrix} \end{aligned}$$

$$\lambda\left[T\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = \lambda\begin{pmatrix} x-2y \\ 2z-x \end{pmatrix}$$

$$
\begin{aligned}
&= \begin{pmatrix} \lambda(x-2y) \\ \lambda(2z-x) \end{pmatrix} \\
&= \begin{pmatrix} (\lambda x) - 2(\lambda y) \\ 2(\lambda z) - (\lambda x) \end{pmatrix};
\end{aligned}
$$

de donde,

$$
T\left[\lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = \lambda \left[T \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right].
$$

□

Ejemplo 2. Sea T: $R^3 \longrightarrow \mathcal{P}_2$, tal que $T \begin{pmatrix} a \\ b \\ c \end{pmatrix} = (a+c)x + 2bx^2$. Veamos que T es una transformación lineal.

Para verificar la *propiedad aditiva*, tomemos dos vectores de R^3, $(a_1, b_1, c_1)^T$ y $(a_2, b_2, c_2)^T$. Entonces,

$$
\begin{aligned}
T\left[\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}\right] &= T \begin{pmatrix} a_1 + a_2 \\ b_1 + b_2 \\ c_1 + c_2 \end{pmatrix} \\
&= [(a_1 + a_2) + (c_1 + c_2)]x + 2(b_1 + b_2)x^2 \\
&= [(a_1 + c_1)x + 2b_1x^2] + [(a_2 + c_2)x + 2b_2x^2] \\
&= T \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + T \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}.
\end{aligned}
$$

Para verificar que T satisface la *propiedad homogénea*, tomemos el escalar λ y el vector $(a, b, c)^T$. Entonces

$$
\begin{aligned}
T\left[\lambda \begin{pmatrix} a \\ b \\ c \end{pmatrix}\right] &= T \begin{pmatrix} \lambda a \\ \lambda b \\ \lambda c \end{pmatrix} \\
&= (\lambda a + \lambda c)x + 2\lambda bx^2 \\
&= \lambda[(a+c)x + 2bx^2] \\
&= \lambda \left[T \begin{pmatrix} a \\ b \\ c \end{pmatrix}\right].
\end{aligned}
$$

□

Ejemplo 3. Sea T: $R^2 \longrightarrow \mathcal{P}_2$, tal que $T \begin{pmatrix} a \\ b \end{pmatrix} = 1 + ax + 2bx^2$. Determinemos si T es una transformación lineal.

Para verificar la *propiedad aditiva*, tomemos dos vectores de R^2, $(a_1, b_1)^T$ y $(a_2, b_2)^T$. Entonces,

$$
T\left[\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}\right] = T \begin{pmatrix} a_1 + a_2 \\ b_1 + b_2 \end{pmatrix} = 1 + (a_1 + a_2)x + 2(b_1 + b_2)x^2,
$$

pero

$$
T \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + T \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = (1 + a_1x + 2b_1x^2) + (1 + a_2x + 2b_2x^2) = 2 + (a_1 + a_2)x + 2(b_1 + b_2)x^2,
$$

de donde,

$$T\left[\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}\right] \neq T\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + T\begin{pmatrix} a_2 \\ b_2 \end{pmatrix};$$

por tanto, T no es una transformación lineal. □

Ejemplo 4. Sea A una matriz $m \times n$ y sea $T : R^n \longrightarrow R^m$ la función tal que $T(\mathbf{x}) = A\mathbf{x}$. Determinemos si T es una transformación lineal.

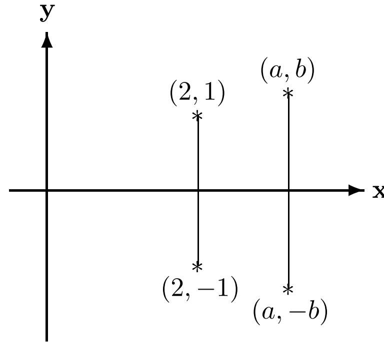
Por el Teorema 2 del Capítulo 2, dados $\mathbf{x}, \mathbf{y} \in R^n$ y $\lambda \in R$, $T(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A\mathbf{x}+A\mathbf{y} = T(\mathbf{x})+T(\mathbf{y})$ y $T(\lambda\mathbf{x}) = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda T(\mathbf{x})$, de donde concluimos que T es una transformación lineal. □

A las transformaciones lineales como la del Ejemplo 4, las llamamos *transformaciones matriciales*. Veremos que todas las transformaciones lineales de R^n en R^m son matriciales. Por ejemplo, notemos que la transformación lineal del Ejemplo 1 es una transformación matricial:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x-2y \\ 2z-x \end{pmatrix} = x\begin{pmatrix} 1 \\ -1 \end{pmatrix} + y\begin{pmatrix} -2 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ -1 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

y por el resultado del Ejemplo 4, esto seria suficiente para demostrar que T es una transformación lineal.

Ejemplo 5. Verifiquemos que la función que a cada punto de R^2 le asigna el punto de su reflexión a través del Eje X es una transformación lineal.



Reflexión en el plano a través del Eje X

De la figura anterior, es fácil ver que la función en cuestión es $T : R^2 \longrightarrow R^2$, donde $T\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ -b \end{pmatrix}$. Por lo tanto,

$$T\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ -b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix},$$

lo que muestra que T es una transformación matricial y por el resultado del Ejemplo 4 una transformación lineal. □

Ejemplo 6. Sea $T : V \longrightarrow W$ la función tal que $T(\mathbf{v}) = \mathbf{0}$ para todo $\mathbf{v} \in V$. Verifiquemos que T es una transformación lineal.

Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ vectores de V. Por ser V un espacio vectorial, $\mathbf{v}_1 + \mathbf{v}_2$ está en V y por tanto, $T(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{0}$. De otro lado, $T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$, de donde concluimos que la propiedad aditiva se satisface. Queda como ejercicio para el lector verificar la propiedad homogénea. A esta transformación lineal la llamamos *transformación nula*. □

Ejemplo 7. Sea $T : V \longrightarrow V$ la función tal que $T(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v} \in V$. Es fácil verificar que T es una transformación lineal, lo que dejamos como ejercicio para el lector. A esta transformación lineal la llamamos *transformación idéntica*. □

Podemos ver que las propiedades que caracterizan a una transformación lineal nos permiten demostrar que la imagen de una combinación lineal de vectores por medio de una transformación lineal es la combinación lineal de las imágenes de los vectores con los mismos coeficientes de la combinación lineal inicial. La demostración consiste básicamente en aplicar la propiedad aditiva iteradamente y luego aplicar la propiedad homogénea a cada sumando.

Teorema 1 [*Transformación lineal de combinaciones lineales*].
Sean $T : V \longrightarrow W$ una transformación lineal, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vectores de V y $\lambda_1, \lambda_2, \ldots, \lambda_n$ escalares de R. Entonces,

$$T(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n) = \lambda_1 T(\mathbf{v}_1) + \lambda_2 T(\mathbf{v}_2) + \ldots + \lambda_n T(\mathbf{v}_n).$$

De este resultado, se tiene trivialmente que una transformación lineal asigna al vector cero del dominio, el vector cero del codominio y, por su importancia, lo enunciamos en el siguiente corolario.

Corolario 1.1 [*Condición necesaria de una transformación lineal*].
Sea $T : V \longrightarrow W$ una transformación lineal, entonces

$$T(\mathbf{0}) = \mathbf{0}.$$

El Teorema 1 también establece que, para el caso de las transformaciones lineales de R^n a R^m, las rectas son enviadas en rectas o puntos y los planos son enviados en planos, rectas o puntos. Adicionalmente, el Teorema 1 permite demostrar que una transformación lineal asigna a un subespacio de dimensión k del dominio, un subespacio del codominio de dimensión igual o menor a k (Ver algunos casos en Ejercicio 8).

Recordemos que dos funciones definidas sobre un mismo dominio y codominio son iguales, si y solo si, tienen las mismas imágenes para todos y cada uno de los elementos del dominio. Aunque una transformación lineal es una función, sus características especiales simplifican enormemente la propiedad de igualdad entre transformaciones lineales, como lo expresamos en el siguiente teorema.

Teorema 2 [*Caracterización de la igualdad de transformaciones lineales*].
Sean $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base del espacio vectorial V y $T : V \longrightarrow W$ y $S : V \longrightarrow W$ dos transformaciones lineales.

$$T = S, \text{ si y solo si, } S(\mathbf{v}_1) = T(\mathbf{v}_1), S(\mathbf{v}_2) = T(\mathbf{v}_2), \ldots, S(\mathbf{v}_n) = T(\mathbf{v}_n).$$

Demostración: Por la igualdad entre funciones, es claro que si $T = S$, las imágenes de los elementos de la base bajo las dos transformaciones son iguales. Para demostrar la otra implicación, recordemos que como $\mathcal{B}$ es una base de V, para cada vector $\mathbf{v}$ de V, existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que $\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n$. Por el Teorema 1 y la igualdad de las imágenes de los elementos de la base bajo las dos transformaciones, tenemos

$$\begin{aligned} T(\mathbf{v}) &= T(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n) \\ &= \lambda_1 T(\mathbf{v}_1) + \lambda_2 T(\mathbf{v}_2) + \ldots + \lambda_n T(\mathbf{v}_n) \\ &= \lambda_1 S(\mathbf{v}_1) + \lambda_2 S(\mathbf{v}_2) + \ldots + \lambda_n S(\mathbf{v}_n) \\ &= S(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n) \\ &= S(\mathbf{v}) \end{aligned}$$

□

Por los teoremas anteriores, es fácil ver que si conocemos la imagen de cada uno de los elementos de una base del dominio de una transformación lineal, podemos conocer la imagen de cualquier otro vector del dominio. En otras palabras, una transformación lineal queda completamente determinada por las imágenes de cada uno de los elementos de una base del dominio, como lo enunciamos en el siguiente teorema.

Teorema 3 [*Determinación de una transformación lineal*].
Si $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base del espacio vectorial V, existe una única transformación lineal
$T : V \longrightarrow W$, tal que $\mathbf{w}_1 = T(\mathbf{v}_1), \mathbf{w}_2 = T(\mathbf{v}_2), \ldots, \mathbf{w}_n = T(\mathbf{v}_n)$, con $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n \in W$.

Demostración: Tenemos que $\mathcal{B}$ es una base de V, así que $\mathcal{B}$ es un conjunto generador de V y por tanto, para cualquier vector $\mathbf{v}$ de V existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que $\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n$. Así, que si sabemos que $\mathbf{w}_1 = T(\mathbf{v}_1), \mathbf{w}_2 = T(\mathbf{v}_2), \ldots, \mathbf{w}_n = T(\mathbf{v}_n)$, podemos encontrar la imagen de cualquier vector $\mathbf{v}$ de V. En efecto, por el Teorema 1,

$$T(\mathbf{v}) = T(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n) = \lambda_1 T(\mathbf{v}_1) + \lambda_2 T(\mathbf{v}_2) + \ldots + \lambda_n T(\mathbf{v}_n) = \lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2 + \ldots + \lambda_n\mathbf{w}_n.$$

Nos queda por demostrar la unicidad de esta transformación. Supongamos que existen dos transformaciones lineales T_1 y T_2 tales que $T_1(\mathbf{v}_i) = \mathbf{w}_i = T_2(\mathbf{v}_i)$ para $i = 1, 2, \ldots, n$. Por el Teorema 2, T_1 y T_2 son la misma transformación lineal. □

Ejemplo 8. Sea $T : \mathcal{P}_2 \longrightarrow R^3$ la transformación lineal tal que $T(1) = (1 \ \ 0 \ \ 0)^T$, $T(x) = (1 \ \ 1 \ \ 0)^T$ y $T(x^2) = (1 \ \ 1 \ \ 1)^T$.

Calculemos $T(a+bx+cx^2)$. Dado que $\{1, x, x^2\}$ es la base canónica de $\mathcal{P}_2$, $a+bx+cx^2 = a\ (1)+b\ (x)+c\ (x^2)$, por lo tanto,

$$T(a + bx + cx^2) = a\ T(1) + b\ T(x) + c\ T(x^2) = a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a+b+c \\ b+c \\ c \end{pmatrix}.$$

□

Ejemplo 9. Dados los vectores $\mathbf{u}_1 = (-1 \ \ 3 \ \ -2)^T$ y $\mathbf{u}_2 = (2 \ \ 0 \ \ 1)^T$de R^3, encontremos una transformación lineal T de R^2 en el plano $H = \left\{\mathbf{v} \in R^3 : \mathbf{v} = t\mathbf{u}_1 + s\mathbf{u}_2, t, s \in R\right\}$ de R^3.

Si tomamos $\{\mathbf{e}_1, \mathbf{e}_2\}$, la base canónica de R^2, y dos vectores arbitrarios de H (no necesariamente diferentes), por ejemplo, $\mathbf{u}_1$ y $\mathbf{u}_2$, podemos definir $T(\mathbf{e}_1) = \mathbf{u}_1$ y $T(\mathbf{e}_2) = \mathbf{u}_2$, de tal manera que

$$T \begin{pmatrix} x \\ y \end{pmatrix} = x\ T(\mathbf{e}_1) + y\ T(\mathbf{e}_2) = x \begin{pmatrix} -1 \\ 3 \\ -2 \end{pmatrix} + y \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -x+2y \\ 3x \\ -2x+y \end{pmatrix}.$$

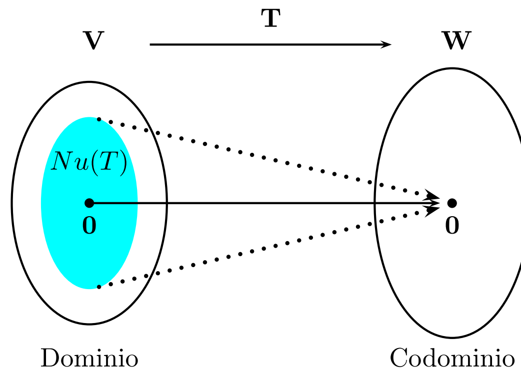
Existe otra transformación lineal de R^2 en el plano H? □

5.3. Espacios Vectoriales Asociados a una Transformación Lineal

Definimos ahora, dos conjuntos asociados a una transformación lineal, los cuales resultan ser Espacios Vectoriales fundamentales para determinar propiedades de este tipo especial de funciones.

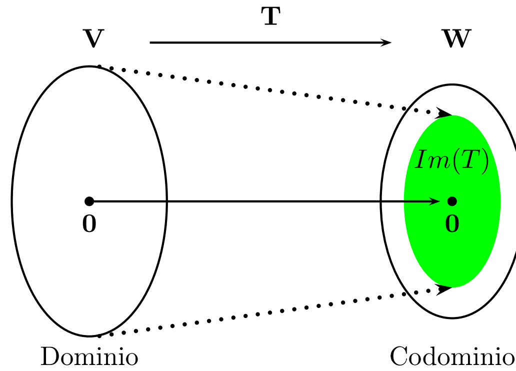
Definición 2 [*Núcleo de una transformación lineal*]. Dada una transformación lineal $T : V \longrightarrow W$, definimos *núcleo de* T como el conjunto $Nu(T)$ de todos los vectores de V cuya imagen es el vector $\mathbf{0}$ de W. En otras palabras,

$$Nu(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}.$$



Definición 3 [*Imagen de una transformación lineal*]. Dada una transformación lineal $T : V \longrightarrow W$, definimos *imagen de* T como el conjunto $Im(T)$ de todos los vectores $\mathbf{w}$ de W para los cuales existe un vector $\mathbf{v}$ de V, tal que $T(\mathbf{v}) = \mathbf{w}$. En otras palabras,

$$Im(T) = \{\mathbf{w} \in W : \text{ existe } \mathbf{v} \in V \text{ tal que } T(\mathbf{v}) = \mathbf{w}\}.$$



Ejemplo 10. Consideremos la transformación lineal $T : \mathcal{M}_{2\times 2} \longrightarrow \mathcal{M}_{2\times 2}$, tal que $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & c+d \end{pmatrix}$ e identifiquemos los conjuntos $Nu(T)$ e $Im(T)$.

Como $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & c+d \end{pmatrix} = \mathbf{0}$ implica que $a = b = 0$ y $c = -d$, concluimos que

$$Nu(T) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a = 0, b = 0, c = -d, \quad a, b, c, d \in R \right\} = \left\{ \begin{pmatrix} 0 & 0 \\ -r & r \end{pmatrix} : r \in R \right\}.$$

Como las únicas matrices que son imágenes bajo T tienen la forma $\begin{pmatrix} r & s \\ 0 & t \end{pmatrix}$ (en efecto, $T\begin{pmatrix} r & s \\ t-\alpha & \alpha \end{pmatrix} = \begin{pmatrix} r & s \\ 0 & t \end{pmatrix}$, para cualquier $\alpha \in R$), tenemos que $Im(T) = \left\{ \begin{pmatrix} r & s \\ 0 & t \end{pmatrix}, \quad r, s, t \in R \right\}$. □

Ejemplo 11. Sea $T : \mathcal{P}_2 \longrightarrow R^3$ la transformación lineal del Ejemplo 8. Identifiquemos $Nu(T)$ e $Im(T)$.

Como vimos en el Ejemplo 8, $T(a+bx+cx^2) = \begin{pmatrix} a+b+c \\ b+c \\ c \end{pmatrix}$, de tal forma que $T(a+bx+cx^2) = \mathbf{0}$ implica que $a = b = c = 0$ y que cualquier vector de R^3 es imagen bajo T (en efecto, $T((p-q)+(q-r)x+rx^2) = (p,q,r)^T$, para cualquier vector $(p,q,r)^T \in R^3$), por lo tanto, $Nu(T) = \{\mathbf{0}\}$ e $Im(T) = R^3$. □

Como ocurrió con los conjuntos asociados a una matriz, los conjuntos asociados a una transformación lineal que acabamos de definir también son espacios vectoriales. En efecto, el núcleo es un subespacio del dominio de la transformación y la imagen es un subespacio del codominio de la transformación, lo cual demostramos en el siguiente teorema.

Teorema 4 [*Subespacios asociados a una transformación lineal*].
Sean V y W espacios vectoriales y $T : V \longrightarrow W$ una transformación lineal. Entonces

1. $Nu(T)$ es subespacio vectorial de V.

2. $Im(T)$ es subespacio vectorial de W.

Demostración: Por el Teorema 1 del Capítulo 4, un subconjunto H no vacío de un espacio vectorial es un subespacio vectorial, si y solo si, los elementos de H satisfacen las propiedades clausurativas para la suma y el producto por escalar (Axiomas 1 y 6 de la definición de espacio vectorial).

1. Por el Corolario 1.1, el vector $\mathbf{0}$ es un elemento de $Nu(T)$, asi que $Nu(T)$ es no vacío. De otro lado, si tomamos dos vectores $\mathbf{u}$ y $\mathbf{v}$ de $Nu(T)$ y un escalar λ, tenemos que $T(\mathbf{u}) = \mathbf{0}$ y $T(\mathbf{v}) = \mathbf{0}$, de modo que

$$\begin{aligned} T(\mathbf{u}+\mathbf{v}) &= T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0} \\ T(\lambda \mathbf{u}) &= \lambda T(\mathbf{u}) = \lambda \mathbf{0} = \mathbf{0}, \end{aligned}$$

de donde concluimos que $\mathbf{u}+\mathbf{v}$ y $\lambda\mathbf{u}$ están en $Nu(T)$.

2. De nuevo, por el Corolario 1.1, el vector $\mathbf{0}$ es un elemento de $Im(T)$, asi que $Im(T)$ es no vacío y si tomamos dos vectores $\mathbf{w}_1$ y $\mathbf{w}_2$ de $Im(T)$ y un escalar λ, tenemos que existen $\mathbf{v}_1$ y $\mathbf{v}_2$, vectores de V tales que $T(\mathbf{v}_1) = \mathbf{w}_1$ y $T(\mathbf{v}_2) = \mathbf{w}_2$, de modo que

$$\begin{aligned} \mathbf{w}_1 + \mathbf{w}_2 &= T(\mathbf{v}_1) + T(\mathbf{v}_2) = T(\mathbf{v}_1 + \mathbf{v}_2) \\ \lambda \mathbf{w}_1 &= \lambda T(\mathbf{v}_1) = T(\lambda \mathbf{v}_1) \end{aligned}$$

de donde concluimos que $\mathbf{w}_1 + \mathbf{w}_2$ y $\lambda\mathbf{w}_1$ están en $Im(T)$. □

5.4. Matriz Asociada a una Transformación Lineal

Ya sabemos que una transformación lineal queda completamente determinada por la forma como la transformación actúa en una base del dominio y que la función de R^n a R^m definida por $T(\mathbf{x}) = A\mathbf{x}$ para una matriz $m \times n$ dada es una transformación lineal, la cual llamamos *transformación matricial.* En esta sección, demostraremos que toda transformación lineal entre espacios vectoriales de dimensión finita puede ser representadaçomo una transformación matricial. Con este propósito, dadas una base del dominio y una base del codominio, para cada transformación lineal, definamos una matriz asociada a ella, como la matriz cuyas columnas son las coordenadas en la base del codominio de las imágenes bajo la transformación de los elementos de la base del dominio.

Definición 4 [*Matriz asociada a una transformación lineal respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$*]. Dadas la transformación lineal $T : V \longrightarrow W$ y las bases $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}'$ de V y W, respectivamente, definimos como *matriz asociada a la transformación lineal T respecto a las bases* $\mathcal{B}$ y $\mathcal{B}'$ a la matriz

$$[A_T]_{\mathcal{B}\mathcal{B}'} = [[T(\mathbf{v}_1)]_{\mathcal{B}'} \; [T(\mathbf{v}_2)]_{\mathcal{B}'} \cdots [T(\mathbf{v}_n)]_{\mathcal{B}'}] .$$

Mientras no haya necesidad de aclarar, denotaremos simplemente por A_T a la matriz asociada a la transformación lineal, respecto a las bases dadas.

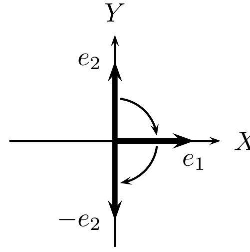
Ejemplo 12. Sea $T : \mathcal{P}_2 \longrightarrow R^2$ la transformación lineal tal que $T(a + bx + cx^2) = \begin{pmatrix} a-b \\ c \end{pmatrix}$ y sean $\mathcal{B} = \{1, 1+x, 1+x-x^2\}$ y $\mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ bases de $\mathcal{P}_2$ y R^2, respectivamente. Encontremos la matriz asociada a la transformación, respecto a las bases dadas.

Como $\mathcal{B}'$ es la base canónica de R^2, las coordenadas de un vector de R^2 coinciden con el vector; por lo tanto, tenemos que $T(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = [T(1)]_{\mathcal{B}'}$, $T(1+x) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = [T(1+x)]_{\mathcal{B}'}$ y $T(1+x-x^2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = [T(1+x-x^2)]_{\mathcal{B}'}$, de donde la matriz asociada a T respecto a las bases dadas es

$$A_T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

□

Ejemplo 13. Calculemos la matriz asociada a la transformación lineal S, que a cada vector del plano cartesiano lo rota 90° en el sentido de las manecillas del reloj, respecto a la base canónica de R^2.

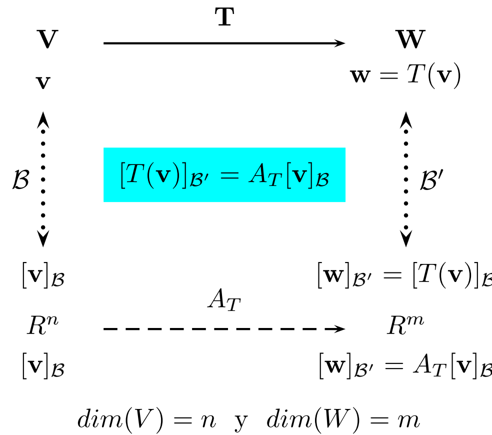


Sea $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$ la base canónica de R^2. Por la definición de S y la figura anterior, $S(\mathbf{e}_1) = -\mathbf{e}_2$ y $S(\mathbf{e}_2) = \mathbf{e}_1$; por lo tanto, la matriz asociada a S respecto a la base canónica es

$$A_S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

□

Podemos ver que la matriz asociada a una transformación lineal permite expresar el vector de coordenadas de la imagen de cualquier vector, en términos de su vector de coordenadas respecto a la base del dominio, de tal manera que, en términos de los vectores de coordenadas, todas las transformaciones lineales resultan ser matriciales, como lo expresa la siguiente propiedad de la matriz asociada a una transformación lineal y lo ilustra el siguiente diagrama.



Teorema 5 [*Caracterización de la matriz asociada a una transformación lineal*].
Dadas la transformación lineal $T : V \longrightarrow W$, con V y W espacios vectoriales de dimensión finita y las bases $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}'$ de V y W, respectivamente, la matriz asociada a la transformación T respecto de estas bases, A_T, es la única matriz tal que, para todo $\mathbf{v} \in V$, $[T(\mathbf{v})]_{\mathcal{B}'} = A_T[\mathbf{v}]_{\mathcal{B}}$.

Demostración: Si $[\mathbf{v}]_{\mathcal{B}} = \left(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n)^T\right)$, entonces $\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n$ y por el Teorema 1, $T(\mathbf{v}) = \lambda_1 T(\mathbf{v}_1) + \lambda_2 T(\mathbf{v}_2) + \ldots + \lambda_n T(\mathbf{v}_n)$. De donde, por el Teorema 14 del Capítulo 4, la combinación lineal se conserva para los vectores de coordenadas respectivos respecto a una misma base; es decir, $[T(\mathbf{v})]_{\mathcal{B}'} = \lambda_1[T(\mathbf{v}_1)]_{\mathcal{B}'} + \lambda_2[T(\mathbf{v}_2)]_{\mathcal{B}'} + \ldots + \lambda_n[T(\mathbf{v}_n)]_{\mathcal{B}'}$. Así que, por definición de $A\mathbf{x}$, tenemos que

$$[T(\mathbf{v})]_{\mathcal{B}'} = A_T[\mathbf{v}]_{\mathcal{B}}.$$

La unicidad se deja como ejercicio para el lector. □

Ejemplo 14. Verifiquemos el teorema anterior con la transformación lineal y las bases del Ejemplo 12.

Si $T : \mathcal{P}_2 \longrightarrow R^2$ es la transformación lineal $T(a + bx + cx^2) = \begin{pmatrix} a - b \\ c \end{pmatrix}$ y $\mathcal{B} = \{1, 1 + x, 1 + x - x^2\}$

y $\mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ son bases de $\mathcal{P}_2$ y R^2, respectivamente, la matriz asociada a la transformación, respecto a las bases dadas es $A_T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. De otro lado, puesto que la base $\mathcal{B}'$ es la base canónica de R^2, $[T(a+bx+cx^2)]_{\mathcal{B}'} = \begin{pmatrix} a-b \\ c \end{pmatrix}$ y, al resolver la ecuación $a+bx+cx^2 = \lambda_1(1)+\lambda_2(1+x)+\lambda_3(1+x-x^2)$, tenemos que $\lambda_1 = a-b$, $\lambda_2 = b+c$ y $\lambda_3 = -c$, asi que $[a+bx+cx^2]_{\mathcal{B}} = \begin{pmatrix} a-b \\ b+c \\ -c \end{pmatrix}$.

Finalmente, comprobamos que $[T(a+bx+cx^2)]_{\mathcal{B}'} = A_T[a+bx+cx^2]_{\mathcal{B}}$, ya que

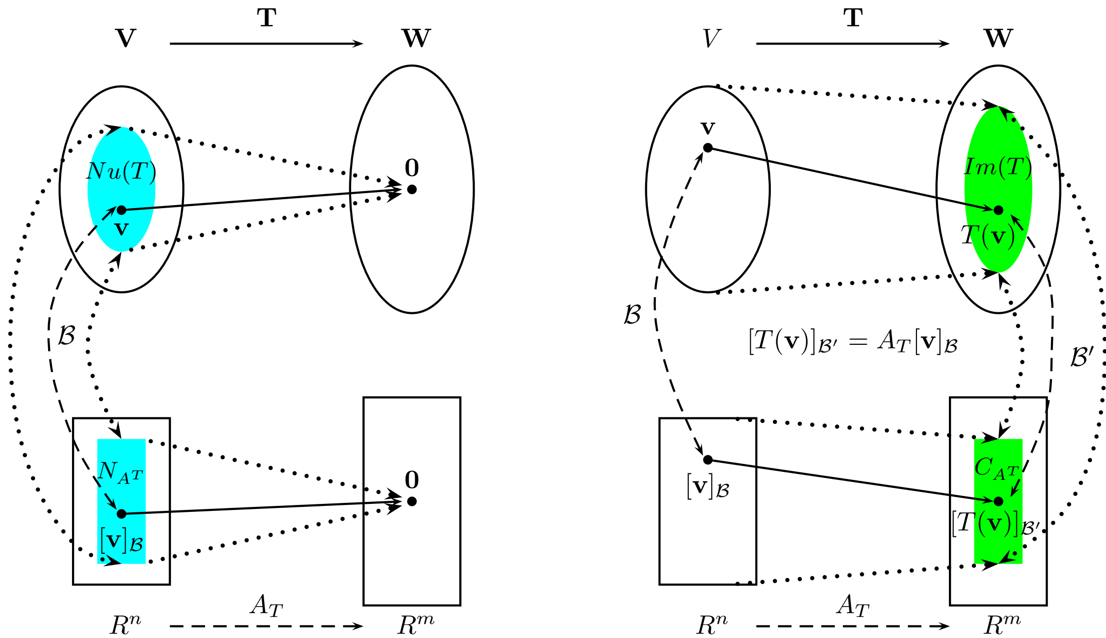
$$\begin{pmatrix} a-b \\ c \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a-b \\ b+c \\ -c \end{pmatrix}.$$

□

El hecho de que, fijadas las bases del dominio y codominio de una transformación lineal, la transformación tenga una matriz asociada, nos permite establecer las siguientes relaciones entre el espacio nulo de la matriz y el núcleo de la transformación y entre el espacio columna de la matriz y la imagen de la transformación.

Teorema 6 [*Equivalencia entre los núcleos y las imágenes de una transformación lineal y los respectivos conjuntos de sus matrices asociadas*]
Dadas la transformación lineal $T : V \longrightarrow W$, con V y W espacios vectoriales de dimensión finita, y las bases $\mathcal{B}$ y $\mathcal{B}'$ de V y W, respectivamente, si A_T es la matriz asociada a la transformación respecto a las bases dadas, entonces

1. $\mathbf{v} \in Nu(T)$, si y solo si, $[\mathbf{v}]_{\mathcal{B}} \in N_{A_T}$
2. $\mathbf{w} \in Im(T)$, si y solo si, $[\mathbf{w}]_{\mathcal{B}'} \in C_{A_T}$



Demostración:

1. Tenemos que

$$\begin{array}{llll} \mathbf{v} \in Nu(T) & \text{si y solo si} & T(\mathbf{v}) = \mathbf{0} \in W & \\ & \text{si y solo si} & [T(\mathbf{v})]_{\mathcal{B}'} = \mathbf{0} & \text{(Teorema 14, Capítulo 4)} \\ & \text{si y solo si} & A_T[\mathbf{v}]_{\mathcal{B}} = \mathbf{0} & \text{(Teorema 5)} \\ & \text{si y solo si} & [\mathbf{v}]_{\mathcal{B}} \in N_{A_T}\ . & \end{array}$$

2. Similarmente,

$$\begin{array}{llll} \mathbf{w} \in Im(T) & \text{si y solo si} & \text{existe } \mathbf{v} \in V \text{ tal que } T(\mathbf{v}) = \mathbf{w} & \\ & \text{si y solo si} & [\mathbf{w}]_{\mathcal{B}'} = [T(\mathbf{v})]_{\mathcal{B}'} = A_T[\mathbf{v}]_{\mathcal{B}} & \text{(Teorema 5)} \\ & \text{si y solo si} & [\mathbf{w}]_{\mathcal{B}'} \in C_{A_T}\ . & \end{array}$$

□

Es importante que resaltemos que $Nu(T)$ es un subespacio de V, el dominio de T, y que N_{A_T} es un subespacio de R^n, de tal manera que si $\mathbf{v} \in Nu(T)$, el vector que está en N_{A_T} es el vector de coordenadas de $\mathbf{v}$ respecto a la base dada de V, no $\mathbf{v}$. Similarmente, $Im(T)$ es un subespacio de W, el codominio de T, y C_{A_T} es un subespacio de R^m y si $\mathbf{w} \in Im(T)$, el vector que está en el espacio columna de la matriz es el vector de coordenadas de $\mathbf{w}$ respecto a la base dada de W, no $\mathbf{w}$. Sin embargo, debemos notar que las dimensiones de $Nu(T)$ e $Im(T)$ coinciden con las de N_{A_T} y C_{A_T} y por consiguiente, si llamamos $\kappa(T) = dim(Nu(T))$, la nulidad de T, y $\tau(T) = dim(Im(T))$, el rango de T, tenemos que

$$\kappa(T) + \tau(T) = dim(V), \tag{5.1}$$

cuya demostración dejamos como ejercicio al lector.

Cuando la transformación lineal va de un espacio en si mismo y tomamos la misma base, tanto en el dominio como en el codominio, hablaremos simplemente de la matriz asociada a la transformación lineal respecto a la base. Así, por ejemplo, cuando tenemos la transformación idéntica, la matriz asociada resulta ser la matriz idéntica, independientemente de la base que se tome (ver Ejercicio 18).

De otro lado, al tomar distintas bases, tenemos distintas matrices asociadas a una misma transformación lineal. El siguiente teorema nos establece una relación entre dos matrices asociadas a una misma transformación lineal y la matriz de transición entre las bases.

Teorema 7 [*Relación entre las matrices asociadas a una transformación lineal*].
Sea $T : V \longrightarrow V$ una transformación lineal, con V un espacio vectorial de dimensión finita (n) y sean $\mathcal{B}$ y $\mathcal{B}'$ dos bases de V. Si P es la matriz de transición de $\mathcal{B}$ a $\mathcal{B}'$, A_T es la matriz asociada a T respecto a $\mathcal{B}$ y A'_T es la matriz asociada a T respecto a $\mathcal{B}'$, entonces

$$A'_T P = P A_T. \tag{5.2}$$

Demostración: Si P es la matriz de transición de $\mathcal{B}$ a $\mathcal{B}'$, $[\mathbf{w}]_{\mathcal{B}'} = P[\mathbf{w}]_{\mathcal{B}}$, para todo $\mathbf{w} \in V$ y, en particular, $[T(\mathbf{v})]_{\mathcal{B}'} = P[T(\mathbf{v})]_{\mathcal{B}}$, para todo $\mathbf{v} \in V$. De otro lado, como A_T es la matriz asociada a T respecto a la base $\mathcal{B}$, $[T(\mathbf{v})]_{\mathcal{B}} = A_T[\mathbf{v}]_{\mathcal{B}}$; así que

$$[T(\mathbf{v})]_{\mathcal{B}'} = P[T(\mathbf{v})]_{\mathcal{B}} = P(A_T[\mathbf{v}]_{\mathcal{B}}) = (PA_T)[\mathbf{v}]_{\mathcal{B}}\ .$$

Por el Teorema 16 del Capítulo 4, P^{-1} es la matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$. Por tanto,

$$[T(\mathbf{v})]_{\mathcal{B}'} = (PA_T)[\mathbf{v}]_{\mathcal{B}} = (PA_T)(P^{-1}[\mathbf{v}]_{\mathcal{B}'}) = (PA_TP^{-1})[\mathbf{v}]_{\mathcal{B}'}.$$

Finalmente, por el Teorema 5, (PA_TP^{-1}) es la matriz asociada a T respecto a la base $\mathcal{B}'$. Como la matriz asociada a T es única,

$$A'_T = PA_TP^{-1},$$

de donde se sigue la ecuación (5.2), que es lo que queríamos demostrar (Ver ilustración en la figura siguiente).

$$V \xrightarrow{\quad T \quad} V$$

$$R^n \overset{A_T}{\dashrightarrow} R^n$$

$$[\mathbf{v}]_{\mathcal{B}} \qquad [T(\mathbf{v}]_{\mathcal{B}} = A_T[\mathbf{v}]_{\mathcal{B}}$$

$P \qquad PA_T \qquad A'_TP \qquad P$

$$[\mathbf{v}]_{\mathcal{B}'} = P[\mathbf{v}]_{\mathcal{B}} \qquad [T\mathbf{v}]_{\mathcal{B}'} = P[T(\mathbf{v})]_{\mathcal{B}} = PA_T[\mathbf{v}]_{\mathcal{B}}$$

$$R^n \overset{A'_T}{\dashrightarrow} R^n$$

$$[\mathbf{v}]_{\mathcal{B}'} \qquad [T\mathbf{v}]_{\mathcal{B}'} = A'_T[(\mathbf{v})]_{\mathcal{B}'} = A'_TP[\mathbf{v}]_{\mathcal{B}}$$

$$dim(V) = n$$

□

Ejemplo 15. Verifiquemos el teorema anterior con la transformación lineal $T : R^2 \longrightarrow R^2$ tal que $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ 2x-y \end{pmatrix}$ y las bases $\mathcal{B} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ y $\mathcal{B}' = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$.

Por ser $\mathcal{B}'$ la base canónica de R^2, la matriz de transición de $\mathcal{B}$ a $\mathcal{B}'$ es $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

De otro lado, tenemos que

$$\begin{array}{lcl} T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} & = & -1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \end{pmatrix}\ ; \text{ es decir, } \left[T\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\ T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} & = & 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 1 \end{pmatrix}\ ; \text{ es decir, } \left[T\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{array}$$

Así que, la matriz asociada a T respecto a $\mathcal{B}$ es $A_T = \left[\ \left[T\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_{\mathcal{B}} \ \left[T\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]_{\mathcal{B}}\ \right] = \begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix}$.
Similarmente, tenemos que

$$\begin{array}{lcl} T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} & = & 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \end{pmatrix}\ ; \text{ es decir, } \left[T\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} & = & 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 1 \end{pmatrix}\ ; \text{ es decir, } \left[T\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{array}$$

Por lo tanto, la matriz asociada a T, respecto a $\mathcal{B}'$ es $A'_T = \left[\ \left[T\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_{\mathcal{B}'} \ \left[T\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]_{\mathcal{B}'}\ \right] = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$.
Finalmente,

$$PA_T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

y

$$A'_TP = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

de donde vemos claramente que se cumple el teorema. □

Todas las matrices asociadas a una misma transformación lineal (pero respecto a diferente base) satisfacen la ecuación (5.2) para alguna matriz invertible (que resulta ser la matriz de transición entre las bases). A las matrices que satisfacen esta ecuación las llamamos *matrices semejantes*, como lo establece la siguiente definición.

Definición 5 [*Matrices semejantes*]. Dadas dos matrices $n \times n$, A y B, decimos que A y B son matrices semejantes, si y solo si, existe una matriz invertible P, tal que $BP = PA$ o lo que es lo mismo, $B = PAP^{-1}$.

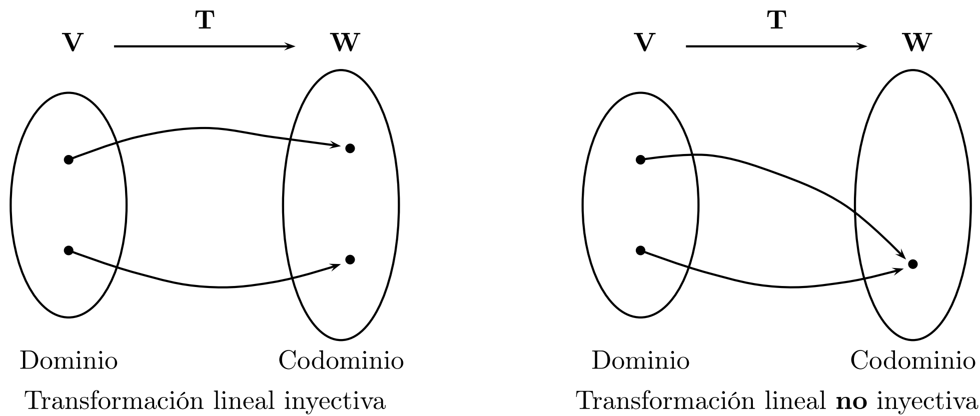
Ejemplo 16. En el Ejemplo 15, las matrices A_T y A'_T son semejantes. □

5.5. Isomorfismos

Aunque hay infinidad de espacios vectoriales, veremos que muchos de ellos son "esencialmente el mismo" respecto a la estructura de espacio vectorial. En esta sección, nos ocuparemos de analizar este concepto de similitud, para lo cual utilizaremos las propiedades de tres tipos de transformaciones lineales especiales, cuyas definiciones damos a continuación.

Definición 6 [*Transformación lineal inyectiva*]. Diremos que $T : V \longrightarrow W$ es una *transformación lineal inyectiva*, si y solo si, para cada $\mathbf{w}$ de $Im(T)$, existe un único $\mathbf{v} \in V$ tal que $T(\mathbf{v}) = \mathbf{w}$.

Una forma de mostrar esta unicidad es suponer que $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ y demostrar que $\mathbf{v}_1 = \mathbf{v}_2$ (ver Ejercicio 19).



Ejemplo 17. Determinemos si la transformación lineal del Ejemplo 10 es inyectiva.

Es fácil ver que la transformación lineal no es inyectiva, ya que, por ejemplo,

$$T\begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 4 \end{pmatrix} = T\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

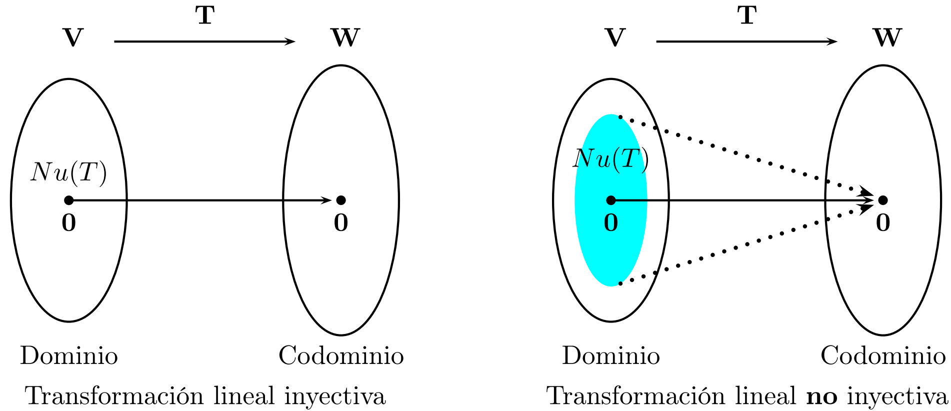
□

Ejemplo 18. Verifiquemos que la transformación lineal del Ejemplo 11 es una transformación lineal inyectiva.

Tenemos que demostrar que si $T(a_1 + b_1x + c_1x^2) = T(a_2 + b_2x + c_2x^2)$, entonces $a_1 = a_2$, $b_1 = b_2$ y $c_1 = c_2$, lo que es fácil de concluir al resolver la ecuación vectorial $\begin{pmatrix} a_1 + b_1 + c_1 \\ b_1 + c_1 \\ c_1 \end{pmatrix} = \begin{pmatrix} a_2 + b_2 + c_2 \\ b_2 + c_2 \\ c_2 \end{pmatrix}$. □

Como lo habiamos mencionado al definir el núcleo y la imagen de una transformación lineal, conocer estos subespacios vectoriales nos permite identificar algunas propiedades de la transformación. Veamos algunas de ellas.

Teorema 8 [*Caracterización de una transformación lineal inyectiva*].
Sea $T : V \longrightarrow W$ una transformación lineal. La transformación T es inyectiva, si y solo si, $Nu(T) = \{\mathbf{0}\}$.



Transformación lineal inyectiva   Transformación lineal **no** inyectiva

Demostración: Supongamos que T es una transformación lineal inyectiva. Como $T(\mathbf{0}) = \mathbf{0}$, por la inyectividad de T, el vector $\mathbf{0}$ tiene que ser el único elemento de $Nu(T)$, por tanto $Nu(T) = \{\mathbf{0}\}$.

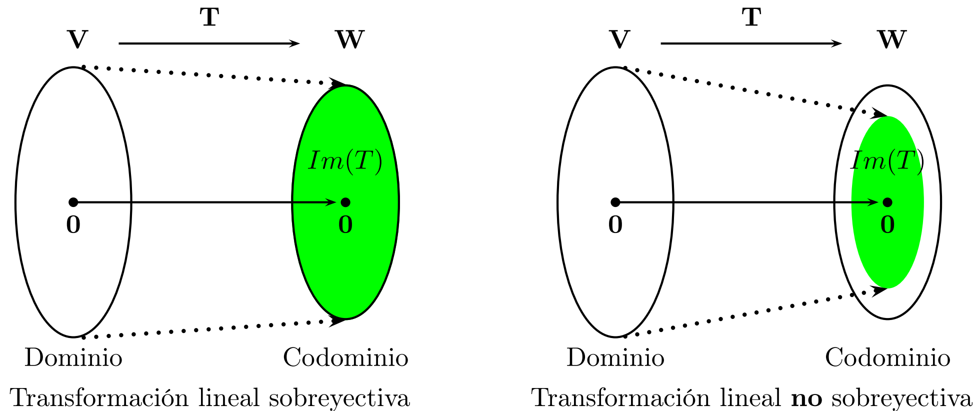
Demostremos ahora la implicación contraria. Supongamos que $Nu(T) = \{\mathbf{0}\}$ y veamos que si $T(\mathbf{u}) = T(\mathbf{v})$, entonces $\mathbf{u} = \mathbf{v}$. Como $T(\mathbf{u}) = T(\mathbf{v})$, entonces $T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u})-T(\mathbf{v}) = \mathbf{0}$, lo que implica que $\mathbf{u}-\mathbf{v} \in Nu(T)$. Pero, como $Nu(T) = \{\mathbf{0}\}$, entonces $\mathbf{u} - \mathbf{v} = \mathbf{0}$ y por tanto $\mathbf{u} = \mathbf{v}$. □

Un resultado bastante útil, que consignaremos en el siguiente teorema, es el que las transformaciones lineales inyectivas envían conjuntos de vectores *l.i.* en conjuntos de vectores *l.i.*; así, conocida una base del dominio, podemos encontrar una base del espacio imagen de la transformación lineal.

Teorema 9 [*Independencia lineal y transformaciones lineales inyectivas*].
Si $T : V \longrightarrow W$ es una transformación lineal inyectiva y $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es un conjunto de vectores *l.i.*, entonces $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}$ es un conjunto de vectores *l.i.*

Demostración: Supongamos que $\lambda_1 T(\mathbf{v}_1)+\lambda_2 T(\mathbf{v}_2)+\ldots+\lambda_n T(\mathbf{v}_n) = \mathbf{0}$. Por el Teorema 1, $T(\lambda_1\mathbf{v}_1+\lambda_2\mathbf{v}_2+\ldots+\lambda_n\mathbf{v}_n) = \mathbf{0}$. Como T es inyectiva, $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$, lo que implica, por la independencia lineal de $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, que $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$, por lo tanto, $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}$ es un conjunto de vectores *l.i.* □

Definición 7 [*Transformación lineal sobreyectiva*]. Diremos que $T : V \longrightarrow W$ es una *transformación lineal sobreyectiva*, si y solo si, $Im(T) = W$.



Transformación lineal sobreyectiva   Transformación lineal **no** sobreyectiva

Ejemplo 19. Determinemos si la transformación lineal del Ejemplo 10 es sobreyectiva.

Dado que $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & c+d \end{pmatrix}$, es fácil ver que para que una matriz esté en la imagen de T, la componente (1,2) debe ser 0, lo que nos indica que $Im(T) \neq \mathcal{M}_{2\times 2}$ y por tanto T no es una transformación lineal sobreyectiva. □

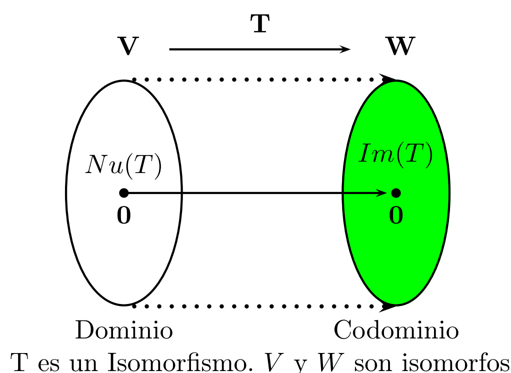
Ejemplo 20. Verifiquemos que la transformación lineal del Ejemplo 12 es sobreyectiva.

Tenemos que verificar que para cualquier vector $\begin{pmatrix} u \\ v \end{pmatrix} \in R^2$, existe un polinomio $(a + bx + cx^2) \in \mathcal{P}_2$, tal que $T(a+bx+cx^2) = \begin{pmatrix} a-b \\ c \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$, para lo cual basta con tomar $c = v$ y, a y b tales que $a - b = u$.
□

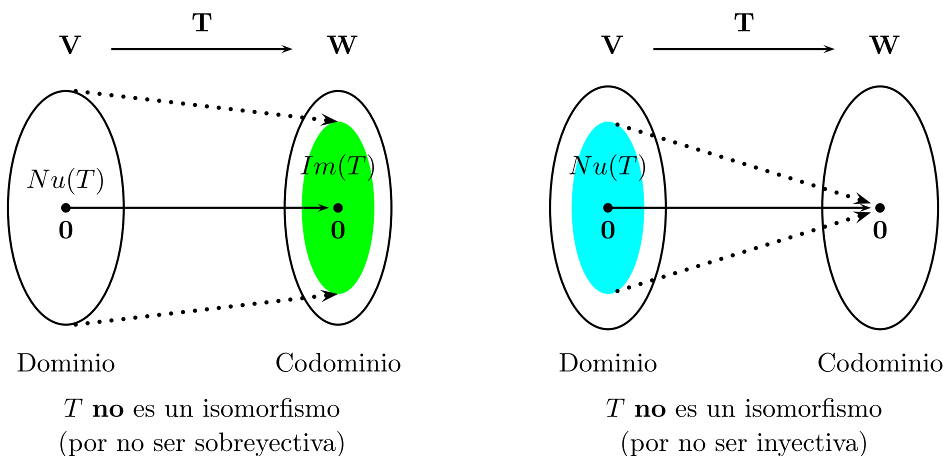
Con los conceptos de inyectividad y sobreyectividad y los resultados hasta ahora obtenidos, podemos ver que si una transformación lineal entre los espacios vectoriales V y W es tanto inyectiva, como sobreyectiva (*isomorfismo*), es porque tienen esencialmente la misma estructura de espacio vectorial; en otra palabras, son "esencialmente los mismos" (*isomorfos*) como lo expresamos en las siguientes definiciones.

Definición 8 [*Isomorfismo*]. Diremos que una transformación lineal $T : V \longrightarrow W$ es un *isomorfismo*, si y solo si, T es una transformación lineal inyectiva y sobreyectiva[1].

Definición 9 [*Espacios vectoriales isomorfos*]. Si dados dos espacios vectoriales V y W, existe un isomorfismo $T : V \longrightarrow W$, diremos que V y W son isomorfos, lo cual denotamos $V \cong W$.



T es un Isomorfismo. V y W son isomorfos



T **no** es un isomorfismo (por no ser sobreyectiva)

T **no** es un isomorfismo (por no ser inyectiva)

[1]Otro nombre que se utiliza para las transformaciones lineales que son inyectivas y sobreyectivas es el de biyectivas.

Ejemplo 21. Determinemos si $T : R^3 \longrightarrow \mathcal{P}_2$ tal que $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a + bx + cx^2$ es un *isomorfismo.*

Debemos determinar si T es inyectiva, y en caso de serlo, determinar si también es sobreyectiva. Para determinar si T es inyectiva supongamos que

$$T\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} = a_1 + b_1x + c_1x^2 = T\begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = a_2 + b_2x + c_2x^2$$

de donde, por igualdad de polinomios, $a_1 = a_2$, $b_1 = b_2$ y $c_1 = c_2$, lo que implica que T es inyectiva.

Ahora, al tomar un polinomio arbitrario $a + bx + cx^2$, podemos mostrar un elemento de R^3, en este caso, $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, tal que $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a + bx + cx^2$ lo que implica que T es sobreyectiva y por lo tanto, T es un isomorfismo y por la Definición 9, podemos concluir que R^3 y $\mathcal{P}_2$ son isomorfos ($R^3 \cong \mathcal{P}_2$). □

Ejemplo 22. La transformación lineal del Ejemplo 10 resultó no ser sobreyectiva (ver Ejemplo 19); por tanto, no es un isomorfismo. □

Cuando el dominio y el codominio tienen la misma dimensión, cualquier transformación lineal que sea inyectiva o sobreyectiva resulta ser un isomorfismo, lo cual demostramos en el siguiente teorema.

Teorema 10 [*Inyectividad y sobreyectividad en espacios de igual dimensión*].
Si $T : V \longrightarrow W$ es una transformación lineal, con V y W espacios vectoriales de dimensión finita y $dim(V) = dim(W)$, entonces

1. si T es inyectiva, T es sobreyectiva.
2. si T es sobreyectiva, T es inyectiva.

Demostración:

1. Si T es inyectiva y $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de V, entonces, $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n\})$ es una base de $Im(T)$, por los Teoremas 1 y 9. Así que $dim(V) = \tau(T) = dim(Im(T))$. Como por hipótesis, $dim(W) = dim(V)$, entonces $dim(W) = dim(Im(T))$; por el Teorema 13 del Capítulo 4 $W = Im(T)$, de donde concluimos que T es sobreyectiva.
2. Si T es sobreyectiva, entonces $\tau(T) = dim(W)$. Como por hipótesis $dimV = dimW$, de la ecuación (5.1), que se desprende del Teorema 6, tenemos que $\kappa(T) = 0$; es decir, $Nu(T) = \{\mathbf{0}\}$, de donde podemos concluir, por el Teorema 8, que T es inyectiva. □

El que dos espacios vectoriales sean isomorfos está intimamente relacionado con la dimensión de ellos, ya que entre dos espacios de igual dimensión siempre se puede construir una transformación lineal tal que a cada vector de una base del dominio le asigna un vector (diferente) de una base del codominio, la cual resulta ser un isomorfismo, como veremos en la demostración del siguiente resultado.

Teorema 11 [*Isomorfismo y dimensión*].
Si V y W son espacios vectoriales de dimensión finita, entonces

$$dim(V) = dim(W), \quad \text{si y solo si,} \quad V \cong W\,.$$

Demostración: Supongamos que $V \cong W$, entonces existe un isomorfismo $T : V \to W$, de tal manera que si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de V, entonces, por ser T una transformación lineal inyectiva, $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n\})$ es una base de $Im(T)$. Además, por ser T una transformación lineal sobreyectiva, $Im(T) = W$, por tanto, $dim(V) = dim(Im(T)) = dim(W)$.

Para mostrar la otra implicación, supongamos que $dim(V) = dim(W)$ y sean $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ y $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ bases de V y W, respectivamente. Podemos ver que la transformación lineal $T : V \rightarrow W$, donde $T(\mathbf{v}_i) = \mathbf{w}_i$ para $i = 1, 2, \ldots, n$ resulta ser un isomorfismo. En efecto, T es sobreyectiva ya que $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ es base, tanto de $Im(T)$, como de W, lo que implica que $Im(T) = W$. En consecuencia, como $dim(V) = dim(W)$, por el Resultado 2 del Teorema 10, T es inyectiva. □

Volviendo sobre la identificación de una transformación lineal y su matriz asociada, podemos ver que las propiedades de las transformaciones lineales se caracterizan por el número de pivotes de la forma escalonada de sus matrices asociadas, como lo establecemos en el siguiente teorema, cuya demostración dejamos como ejercicio para el lector.

Teorema 12 [*Matriz asociada, inyectividad y sobreyectividad*].
Sean $T : V \longrightarrow W$ una transformación lineal, con V y W espacios vectoriales de dimensión finita y A_T la matriz $m \times n$ asociada a la transformación, respecto a dos bases dadas. Entonces,

1. T es inyectiva, si y solo si, la forma escalonada de A_T tiene n pivotes.
2. T es sobreyectiva, si y solo si, la forma escalonada de A_T tiene m pivotes.
3. T es un isomorfismo, si y solo si, A_T es invertible.

5.6. Algebra de transformaciones lineales

Dado que las transformaciones lineales son casos particulares de funciones, podemos definir la suma, la multiplicación por escalar y la composición de transformaciones lineales, como se hace en las funciones de valor real. Adicionalmente, veremos que estas operaciones están relacionadas con las operaciones matriciales de sus matrices asociadas.

Definición 10 [*Suma y producto por escalar de transformaciones lineales*]. Dadas dos transformaciones lineales $T, S : V \longrightarrow W$ y un escalar $\lambda \in R$, definimos $T + S$, *la suma* de T y S, como la función

$$(T+S) : V \longrightarrow W, \quad \text{tal que} \quad (T+S)(\mathbf{v}) = T(\mathbf{v}) + S(\mathbf{v}), \quad \text{para todo} \quad \mathbf{v} \in V$$

y definimos λT, la *multiplicación de T por el escalar λ*, como la función

$$(\lambda T) : V \longrightarrow W, \quad \text{tal que} \quad (\lambda T)(\mathbf{v}) = \lambda[T(\mathbf{v})], \quad \text{para todo} \quad \mathbf{v} \in V.$$

Ejemplo 23. Dadas las transformaciones lineales $T, S : R^2 \longrightarrow \mathcal{P}_1$, tal que $T\begin{pmatrix} a \\ b \end{pmatrix} = a + bx$ y $S\begin{pmatrix} a \\ b \end{pmatrix} = a - 2bx$, calculemos $(T+S)\begin{pmatrix} -2 \\ 5 \end{pmatrix}$ y $(-3T)\begin{pmatrix} -2 \\ 5 \end{pmatrix}$.

Por la definición anterior,

$$(T+S)\begin{pmatrix} -2 \\ 5 \end{pmatrix} = T\begin{pmatrix} -2 \\ 5 \end{pmatrix} + S\begin{pmatrix} -2 \\ 5 \end{pmatrix} = (-2+5x) + (-2-10x) = -4 - 5x$$

y

$$(-3T)\begin{pmatrix} -2 \\ 5 \end{pmatrix} = -3\left[T\begin{pmatrix} -2 \\ 5 \end{pmatrix}\right] = -3(-2+5x) = 6 - 15x.$$

□

Teorema 13 [*Propiedad clausurativa de la suma y el producto por escalar de transformaciones lineales*].
Si $S, T : V \longrightarrow W$ son transformaciones lineales y μ es un escalar, entonces las funciones

$$(T+S) : V \longrightarrow W \ \text{ y } \ (\mu T) : V \longrightarrow W$$

son transformaciones lineales.

Demostración: Sean $\mathbf{v}_1, \mathbf{v}_2 \in V$, entonces

$$\begin{aligned}(S+T)(\mathbf{v}_1+\mathbf{v}_2) &= S(\mathbf{v}_1+\mathbf{v}_2)+T(\mathbf{v}_1+\mathbf{v}_2) = S(\mathbf{v}_1)+S(\mathbf{v}_2)+T(\mathbf{v}_1)+T(\mathbf{v}_2)\\ &= S(\mathbf{v}_1)+T(\mathbf{v}_1)+S(\mathbf{v}_2)+T(\mathbf{v}_2) = (S+T)(\mathbf{v}_1)+(S+T)(\mathbf{v}_2)\end{aligned}$$

$$\begin{aligned}(S+T)(\lambda\mathbf{v}) &= S(\lambda\mathbf{v})+T(\lambda\mathbf{v}) = \lambda\,[S(\mathbf{v})]+\lambda\,[T(\mathbf{v})]\\ &= \lambda[S(\mathbf{v})+T(\mathbf{v})] = \lambda(S+T)(\mathbf{v}),\end{aligned}$$

de donde podemos concluir que $(S+T)$ es también una transformación lineal. De otro lado,

$$\begin{aligned}(\mu T)(\mathbf{v}_1+\mathbf{v}_2) &= \mu[T(\mathbf{v}_1+\mathbf{v}_2)] = \mu[T(\mathbf{v}_1)+T(\mathbf{v}_2)]\\ &= \mu[T(\mathbf{v}_1)]+\mu[T(\mathbf{v}_2)] = (\mu T)(\mathbf{v}_1)+(\mu T)(\mathbf{v}_2)\end{aligned}$$

$$\begin{aligned}(\mu T)(\lambda\mathbf{v}) &= \mu[T(\lambda\mathbf{v})] = \mu[\lambda T(\mathbf{v})]\\ &= (\mu\lambda)[T(\mathbf{v})] = \lambda[(\mu T)(\mathbf{v})],\end{aligned}$$

de donde concluimos que (μT) también es una transformación lineal. □

El conjunto de las transformaciones lineales de un espacio vectorial V en otro espacio vectorial W, con la suma y multiplicación por escalar antes definidas, satisfacen las propiedades que caracterizan a los espacios vectoriales. El Teorema 13 establece las propiedades clausurativas y, en el siguiente teorema, planteamos el resto de las 10 propiedades que demuestran que $\mathcal{L}(V,W)$, el conjunto de las transformaciones lineales de V en W, es un espacio vectorial[2].

Teorema 14 [*Espacio vectorial de transformaciones lineales*].
Dadas las transformaciones lineales $R, S, T : V \longrightarrow W$ y los escalares λ, μ, entonces

1. $R+(S+T) = (R+S)+T$
2. $S+T = T+S$
3. $T+\mathbf{0} = \mathbf{0}+T = T$
4. $T+(-T) = (-T)+T = \mathbf{0}$
5. $\lambda(S+T) = \lambda S+\lambda T$
6. $(\lambda+\mu)T = \lambda T+\mu T$
7. $(\lambda\mu)T = \lambda(\mu T) = \mu(\lambda T)$
8. $1T = T$
9. $0T = \mathbf{0}$

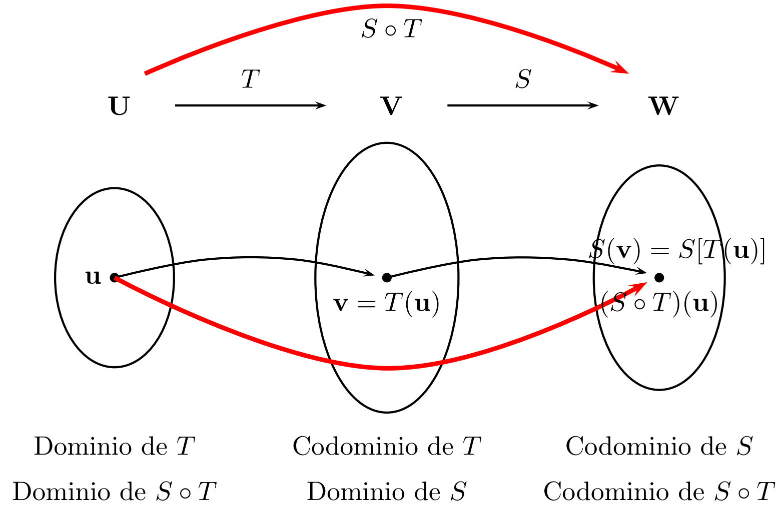
Demostración: Se deja como ejercicio para el lector.

Definamos ahora la composición, una operación que está definida en funciones tales que el codominio de una de las funciones coincide con el dominio de la otra. En este caso, nos restringiremos a transformaciones lineales.

Definición 11 [*Composición de transformaciones lineales*]. Sean $T : U \longrightarrow V$ y $S : V \longrightarrow W$ transformaciones lineales. Definimos $S \circ T$, la *composición de S con T*, como la función

$$(S\circ T) : U \longrightarrow W \quad \text{tal que} \quad (S\circ T)(\mathbf{u}) = S[T(\mathbf{u})] \quad \text{para todo} \quad \mathbf{u} \in U\,.$$

[2]Si tenemos en cuenta que una transformación lineal es una función, por el Teorema 13, podemos concluir que $\mathcal{L}(V,W)$, el conjunto de las transformaciones lineales de V en W, es un subespacio del espacio de funciones de V en W, y por lo tanto, es un espacio vectorial; en consecuencia satisface las propiedades planteadas en el Teorema 14



Teorema 15 [*Propiedad "clausurativa" de la composición de transformaciones lineales*].
Dadas $T : U \longrightarrow V$ y $S : V \longrightarrow W$, transformaciones lineales, la composición de S con T, $(S \circ T) : U \longrightarrow W$, es también una transformación lineal.

Demostración: Sean $\mathbf{u}, \mathbf{u}_1, \mathbf{u}_2 \in U$ y $\lambda \in R$, entonces

$$\begin{aligned}(S \circ T)(\mathbf{u}_1 + \mathbf{u}_2) &= S[T(\mathbf{u}_1 + \mathbf{u}_2)] = S[T(\mathbf{u}_1) + T(\mathbf{u}_2)] \\ &= S[T(\mathbf{u}_1)] + S[T(\mathbf{u}_2)] = (S \circ T)(\mathbf{u}_1) + (S \circ T)(\mathbf{u}_2).\end{aligned}$$

$$\begin{aligned}(S \circ T)(\lambda \mathbf{u}) &= S[T(\lambda \mathbf{u})] = S(\lambda T(\mathbf{u})) \\ &= \lambda[S[T(\mathbf{u})]] = \lambda(S \circ T)(\mathbf{u}).\end{aligned}$$

Por consiguiente, $S \circ T$ es una transformación lineal. $\square$

Ejemplo 24. Dadas las transformaciones lineales $T : R^3 \longrightarrow \mathcal{P}_1$ y $S : \mathcal{P}_1 \longrightarrow R^2$, tal que $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = (a + b) + cx$ y $S(\alpha + \beta x) = \begin{pmatrix} \alpha - \beta \\ \beta \end{pmatrix}$, calculemos $(S \circ T)\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$.

Por la definición,

$$(S \circ T)\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = S\left[T\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}\right] = S\,(-1 + 3x) = \begin{pmatrix} -1 - 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \end{pmatrix}.$$

Observemos que, en general, la composición de transformaciones lineales no es conmutativa. Es más, en muchos casos $S \circ T$ está bien definida pero $T \circ S$ no, como ocurre en el Ejemplo 24, donde $S \circ T$ está bien definida pero $T \circ S$, no. En otros casos, aún estando ambas composiciones bien definidas, sus dominios y/o imágenes pertenecen a espacios vectoriales distintos. $\square$

La composición de transformaciones lineales también satisface propiedades algebraicas, las cuales consignamos en el siguiente teorema.

Teorema 16 [*Propiedades de la composición de transformaciones lineales*].
Sean R, S y T transformaciones lineales tales que, en cada caso, las composiciones están bien definidas y sea λ un escalar, entonces

1. $(T \circ S) \circ R = T \circ (S \circ R)$.
2. $T \circ (S + R) = (T \circ S) + (T \circ R)$.
3. $(T + S) \circ R = (T \circ R) + (S \circ R)$.
4. $\lambda(T \circ S) = (\lambda T) \circ S = T \circ (\lambda S)$.
5. $I \circ T = T \circ I = T$.
6. $\mathbf{0} \circ T = \mathbf{0}$ y $T \circ \mathbf{0} = \mathbf{0}$.

Demostración: Demostremos las Propiedades 1 y 4, las demás las dejamos como ejercicio para el lector.

1. Por definición de la composición de transformaciones lineales,

$$\begin{aligned} [(T \circ S) \circ R](\mathbf{v}) &= (T \circ S)[R(\mathbf{v})] = T[S(R(\mathbf{v}))] \\ &= T[(S \circ R)(\mathbf{v})] = [T \circ (S \circ R)](\mathbf{v}). \end{aligned}$$

4. Igualmente, por la definición de la composición de transformaciones lineales,

$$\begin{aligned} \lambda(T \circ S)(\mathbf{v}) &= \lambda[T(S(\mathbf{v})] = (\lambda T)[S(\mathbf{v})] \\ &= [(\lambda T) \circ S](\mathbf{v}) \end{aligned}$$

de otro lado,

$$\begin{aligned} \lambda(T \circ S)(\mathbf{v}) &= (T \circ S)(\lambda \mathbf{v}) = T(S(\lambda \mathbf{v})) \\ &= T(\lambda S(\mathbf{v})) = [T \circ (\lambda S)](\mathbf{v}). \end{aligned}$$

□

En la Sección 5.2, vimos que existe una estrecha relación entre las transformaciones lineales y las matrices, a saber, si T es una transformación lineal $T : V \longrightarrow W$, $\mathcal{B}$ y $\mathcal{B}'$ son bases de V y W, respectivamente y A_T es la matriz asociada a T, respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$, entonces

$$[T(\mathbf{v})]_{\mathcal{B}'} = A_T[\mathbf{v}]_{\mathcal{B}}.$$

Mediante esta relación entre transformaciones lineales y matrices, podemos establecer una correspondencia entre las operaciones algebraicas de las transformaciones lineales y las operaciones algebraicas de las matrices asociadas a ellas, como lo consignamos en el siguiente teorema.

Teorema 17 [*Transformaciones lineales y matrices asociadas*].
Sean $\mathcal{B}$, $\mathcal{B}'$ y $\mathcal{B}''$ bases de los espacios vectoriales U, V y W, respectivamente; $T, S : U \longrightarrow V$ y $R : V \longrightarrow W$ transformaciones lineales; A_T y A_S las matrices asociadas a T y S respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ y A_R la matriz asociada a R respecto a las bases $\mathcal{B}'$ y $\mathcal{B}''$, entonces la matriz asociada a la transformación lineal

1. $T + S$ respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ es $A_T + A_S$.
2. λT respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ es λA_T.
3. $-T$ respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ es $-A_T$.
4. $T - S$ respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ es $A_T - A_S$.
5. $R \circ T$ respecto a las bases $\mathcal{B}$ y $\mathcal{B}''$ es $A_R A_T$.

Demostración: Demostremos la Propiedad 5; las demás las dejamos como ejercicio para el lector.

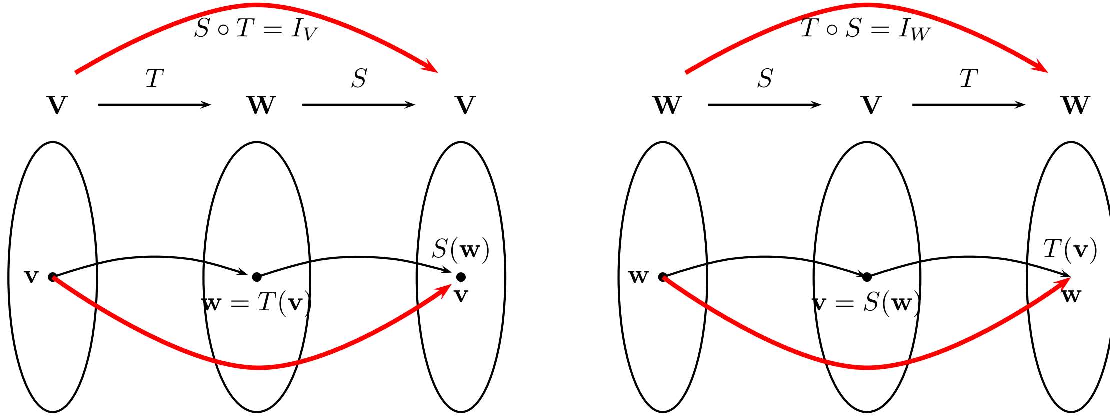
Por la definición de matriz asociada a una transformación lineal,

$$[(R \circ T)(\mathbf{u})]_{\mathcal{B}''} = [R(T(\mathbf{u}))]_{\mathcal{B}''} = A_R[T(\mathbf{u})]_{\mathcal{B}'} = A_R(A_T[\mathbf{u}]_{\mathcal{B}}) = (A_R A_T)[\mathbf{u}]_{\mathcal{B}}$$

y por la propiedad de unicidad de la matriz asociada a $R \circ T$ (Teorema 5), la matriz asociada a $(R \circ T)$ respecto a las bases $\mathcal{B}$ y $\mathcal{B}''$ es $A_R A_T$. □

Como ocurre en general para el conjunto de funciones, la composición de transformaciones lineales permite caracterizar las transformaciones lineales que son invertibles, como lo consignamos en la siguiente definición.

Definición 12 [*Transformación lineal invertible*]. Decimos que $T : V \longrightarrow W$ es una *transformación lineal invertible*, si y sólo si, existe una transformación lineal $S : W \longrightarrow V$ tal que $(T \circ S) = I_W : W \longrightarrow W$ y $(S \circ T) = I_V : V \longrightarrow V$. A la transformación S la llamamos *inversa* de T.



Ejemplo 25. Verifiquemos que $T : R^2 \longrightarrow \mathcal{P}_1$ tal que $T\begin{pmatrix} a \\ b \end{pmatrix} = (a+b) + (a-b)x$ es una transformación lineal invertible.

Veamos que $S : \mathcal{P}_1 \longrightarrow R^2$ tal que $S(u+vx) = \left(\frac{u+v}{2}, \frac{u-v}{2}\right)^T$ es una transformación lineal tal que $T \circ S = I_{\mathcal{P}_1}$ y $S \circ T = I_{R^2}$.

$$(T \circ S)(u+vx) = T\left[S(u+vx)\right] = T\begin{pmatrix} \frac{u+v}{2} \\ \frac{u-v}{2} \end{pmatrix} = \frac{u+v}{2} + \frac{u-v}{2} + \left(\frac{u+v}{2} - \frac{u-v}{2}\right)x = u+vx = I_{\mathcal{P}_1}(u+vx)$$

$$(S \circ T)\begin{pmatrix} a \\ b \end{pmatrix} = S\left[T\begin{pmatrix} a \\ b \end{pmatrix}\right] = S((a+b)+(a-b)x) = \begin{pmatrix} \frac{(a+b)+(a-b)}{2} \\ \frac{(a+b)-(a-b)}{2} \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} = I_{R^2}\begin{pmatrix} a \\ b \end{pmatrix}.$$

En consecuencia, S es la inversa de T; por lo tanto, T es una transformación lineal invertible. □

Al igual que con las funciones en general y que con las matrices, la inversa de una transformación lineal es única, como lo planteamos en el siguiente teorema, cuya demostración dejamos como ejercicio para el lector.

Teorema 18 [*Unicidad de la transformación lineal inversa*].
Sea $T : V \to W$ una transformación lineal. Si existen $S_1, S_2 : W \to V$, dos transformaciones lineales tales que $T \circ S_1 = T \circ S_2 = I_W$ y $S_1 \circ T = S_2 \circ T = I_V$, entonces $S_1 = S_2$.

Similarmente como hicimos en el caso de las matrices, la inversa de la transformación lineal T la denotamos T^{-1}. A continuación, demostramos la relación de equivalencia entre las transformaciones lineales invertibles

y los isomorfismos, que, en particular, nos permite determinar fácilmente cuando una transformación lineal no es invertible.

Teorema 19 [*Transformaciones lineales invertibles e isomorfismos*].
Sea $T : V \to W$ una transformación lineal, entonces T es invertible, si y solo si, T es un isomorfismo.

Demostración: Supongamos que T es invertible y que S es su inversa. Tomemos $\mathbf{v}_1, \mathbf{v}_2 \in V$ y supongamos que $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, así que $S(T(\mathbf{v}_1)) = S(T(\mathbf{v}_2))$, de donde concluimos que $\mathbf{v}_1 = \mathbf{v}_2$ y por consiguiente que T es inyectiva. Tomemos ahora $\mathbf{w} \in W$ y sea $\mathbf{v} = S(\mathbf{w})$, así que $T(\mathbf{v}) = T(S(\mathbf{w})) = \mathbf{w}$, de donde concluimos que T es sobreyectiva y por tanto, es un isomorfismo.

Supongamos ahora que T es un isomorfismo, asi que T es inyectiva y sobreyectiva. Sea $\mathbf{w} \in W$, entonces existe un único $\mathbf{v} \in V$, tal que $T(\mathbf{v}) = \mathbf{w}$. Definimos S, de tal forma que $S(\mathbf{w}) = \mathbf{v}$, cuando $T(\mathbf{v}) = \mathbf{w}$. Es fácil verificar que S es la inversa de T; por tanto, T es invertible. □

Ejemplo 26. Utilicemos el teorema anterior para determinar si la transformación del Ejemplo 1 es invertible.

Si la transformación del Ejemplo 1, que va de R^3 a R^2 fuese invertible, por el teorema anterior, T seria un isomorfismo y por lo tanto, los espacios vectoriales R^3 y R^2 serian isomorfos, lo cual, por el Teorema 11, no es posible ya que sus dimensiones son distintas. En consecuencia, la transformación del Ejemplo 1 no es invertible. □

Por el Teorema 19, podemos concluir que la transformación lineal del Ejemplo 25 es un isomorfismo, pero en general no es fácil determinar que una transformación lineal es un isomorfismo encontrando su transformación inversa. Afortunadamente, podemos demostrar que para determinar si una transformación lineal es invertible es suficiente determinar que una matriz asociada a ella lo sea, dada la relación entre las transformaciones lineales invertibles y las matrices invertibles, que establecemos en el siguiente teorema.

Teorema 20 [*Transformación lineal invertible y matriz asociada*].
Sean $\mathcal{B}$ y $\mathcal{B}'$ bases de los espacios vectoriales V y W, respectivamente, $T : V \to W$ una transformación lineal y A_T la matriz asociada a la transformación respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$. Entonces

1. T es invertible, si y solo si, A_T es invertible.
2. Si T es invertible, entonces A_T^{-1} es la matriz asociada a la transformación lineal T^{-1}, respecto a las bases $\mathcal{B}'$ y $\mathcal{B}$.

Demostración: La dejamos como ejercicio para el lector.

Ejemplo 27. Utilicemos el teorema anterior para determinar si la transformación lineal del Ejemplo 25 es invertible, sin calcular su transformación lineal inversa.

Sean $\mathcal{B}$ y $\mathcal{B}'$ las bases canónicas de R^2 y $\mathcal{P}_1$, respectivamente. La matriz asociada a T en las bases $\mathcal{B}$ y $\mathcal{B}'$ es $A_T = [\ [T(\mathbf{e}_1)]_{\mathcal{B}'}\ [T(\mathbf{e}_1)]_{\mathcal{B}'}\ ] = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, cuyo determinante es -2, por lo tanto A_T es invertible y en consecuencia, por el teorema anterior, T es invertible. □

5.7. Ejercicios

1. Determine cuáles de las siguientes funciones son transformaciones lineales.
   a) $T : R^3 \longrightarrow R^2$ tal que $T(a\ \ b\ \ c)^T = (2a\ \ b - 3c)^T$
   b) $T : R^2 \longrightarrow \mathcal{P}_1$ tal que $T(a\ \ b)^T = (1 + a) - 2bx$
   c) $T : R \longrightarrow \mathcal{P}_2$ tal que $T(a) = a + 2ax + 3ax^2$
   d) $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ tal que $T(a + bx + cx^2) = a + x - (b + c)x^2$
   e) $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_1$ tal que $T(p(x)) = p'(x)$

f) $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_3$ tal que $T(p(x)) = xp(x)$

g) $T : \mathcal{M}_{2\times 2} \longrightarrow \mathcal{P}_2$ tal que $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + bx + cx^2$

h) $T : R^3 \longrightarrow R^2$ tal que $T(a \;\; b \;\; c)^T = (a \;\; 2b - 3c)^T$

i) $T : R^2 \longrightarrow R^3$ tal que $T(a \;\; b)^T = (-3a \;\; b + a \;\; b)^T$

j) $T : \mathcal{M}_{2\times 2} \longrightarrow \mathcal{M}_{2\times 2}$ tal que $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a - c & 0 \\ 0 & d + 2 \end{pmatrix}$

k) $T : R^3 \longrightarrow \mathcal{P}_2$ tal que $T(a \;\; b \;\; c)^T = a + 2bx - 3cx^2$

l) $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ tal que $T(a + bx + cx^2) = a - cx^2$

m) La reflexión a través del plano XY en R^3.

n) $T : R^2 \longrightarrow R^2$ tal que $T(x \;\; y)^T = (\alpha x \;\; y)^T$ (*compresión* $[|\alpha| < 1]$ ó *expansión* $[|\alpha| > 1]$ a lo largo del Eje X)

ñ) La proyección sobre el plano XZ en R^3

o) $T : \mathcal{M}_{3\times 2} \longrightarrow \mathcal{M}_{2\times 3}$ tal que $T(A) = A^T$

2. Demuestre la propiedad homogénea de la transformación nula (Ejemplo 6).

3. Demuestre las propiedades aditiva y homogénea de la transformación idéntica (Ejemplo 7).

4. Considere la transformación matricial de R^2 en R^2, donde la matriz es una matriz elemental. Interprete geométricamente dicha transformación. Qué puede decirse de este tipo de transformaciones en R^n?

5. Demuestre el Teorema 1. *(**AYUDA:** Use repetidamente la propiedad aditiva de la transformación lineal y luego la propiedad homogénea en cada sumando.)*

6. Demuestre el Corolario 1.1.

7. Para cada una de las siguientes transformaciones lineales:

   a) $T : R^3 \longrightarrow R^2$ tal que $T(x \;\; y \;\; z)^T = (x - y \;\; x - z)^T$

   b) $T : R^3 \longrightarrow R^3$ tal que $T(x \;\; y \;\; z)^T = (x - y + z \;\; 0 \;\; 0)^T$

   c) $T : R^2 \longrightarrow R^3$ tal que $T(x \;\; y \;\; z)^T = (2x \;\; 0 \;\; 0)^T$

   d) $T : R^3 \longrightarrow R^3$ tal que $T(x \;\; y \;\; z)^T = (x - y \;\; x \;\; z)^T$

   e) $T : R^3 \longrightarrow R^3$ tal que $T(x \;\; y \;\; z)^T = (2y \;\; 0 \;\; x)^T$

   encuentre la imagen de los siguientes conjuntos, identificándolas geométricamente, si es posible.
   S_1 : La recta que pasa por $P(-1 \;\; 2 \;\; 5)^T$ y $Q(-2 \;\; 0 \;\; 3)^T$.
   S_2 : La recta que pasa por $P(-1 \;\; 2 \;\; 5)^T$ y $Q(2 \;\; -4 \;\; -10)^T$.
   S_3 : El plano que pasa por el origen y tiene vectores directores $\mathbf{u} = (2 \;\; 0 \;\; -1)^T$ y $\mathbf{v} = (0 \;\; -1 \;\; -3)^T$.
   S_4 : El plano que pasa por el punto $Q(-2 \;\; 0 \;\; 3)^T$ y tiene vectores directores $\mathbf{u} = (2 \;\; 0 \;\; -1)^T$ y $\mathbf{v} = (0 \;\; -1 \;\; -3)^T$.

8. Demuestre que una transformación lineal de R^n en R^m envía rectas en rectas o puntos, y a planos en planos, rectas o puntos. Que pasa con los hiperplanos?

9. Cuantas transformaciones lineales de R^2 en el plano $H = \{\mathbf{v} \in R^3 : \mathbf{v} = t\mathbf{u}_1 + s\mathbf{u}_2, t, s \in R\}$ de R^3 existen? (Ver Ejemplo 9).

10. Encuentre una matriz asociada a la transformación lineal de R^3 que rota todo vector 45° alrededor del Eje Y.

11. Determine cuáles de las transformaciones lineales del Ejercicio 1 son transformaciones matriciales.

12. Dada la transformación lineal $T : V \longrightarrow W$ y las bases $\mathcal{B}$ y $\mathcal{B}'$ de V y W, respectivamente, demuestre que la matriz A_T, tal que $[T(\mathbf{v})]_{\mathcal{B}'} = A_T[\mathbf{v}]_{\mathcal{B}}$ para todo $\mathbf{v} \in V$, es única.(***AYUDA:*** *Suponga que existe otra matriz* B *y tome* $\mathbf{v} = \mathbf{v}_i$*, con* $\mathbf{v}_i \in \mathcal{B}$ *para encontrar las columnas de* A_T *y* B).

13. Demuestre el Teorema 12. *(**AYUDA:** Utilice las propiedades de las matrices escalonadas estudiadas en el Capitulo 1 y los resultados del Teorema 6.)*

14. Encuentre una matriz asociada a cada una de las transformaciones lineales del Ejercicio 1, explicitando la bases del dominio y del codominio. Existe otra matriz asociada a cada transformación?

15. Encuentre el núcleo y la imagen de cada una de las transformaciones lineales del Ejercicio 1 y diga cuáles son sus dimensiones.

16. Demuestre que si $T : V \longrightarrow V$ es una transformación lineal, entonces $\kappa(T) + \tau(T) = dim(V)$.

17. Demuestre que si V es un espacio vectorial de dimensión n y $T : V \longrightarrow V$ es la transformación idéntica ($T(\mathbf{v}) = \mathbf{v}$, para todo $\mathbf{v} \in V$), la matriz asociada a T es la matriz identidad de tamaño n, independientemente de la base de V que se tome.

18. Demuestre que $T : V \longrightarrow W$ es una transformación lineal inyectiva (para cada $\mathbf{w} \in Im(T)$, existe un **único** $\mathbf{v} \in V$ tal que $T(\mathbf{v}) = \mathbf{w}$), si y solo si, $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ implica que $\mathbf{v}_1 = \mathbf{v}_2$.

19. Determine si las transformaciones lineales del Ejercicio 1 son inyectivas, sobreyectivas, isomorfismos y/o invertibles.

20. Demuestre el Teorema 14.

21. Encuentre una transformación lineal no trivial entre cada uno de los espacios vectoriales dados y determine si es inyectiva, sobreyectiva, isomorfismo y/o invertible.

    a) $V = R^2$ y $W =$ el plano de R^3 que pasa por el origen y tiene vectores directores $\mathbf{d}_1 = (-1 \;\; 2 \;\; 5)^T$ y $\mathbf{d}_2 = (3 \;\; 0 \;\; -2)^T$

    b) $V = \{p(x) = a + bx + cx^2 : a = 0\}$ y $W = \{p(x) = a + bx + cx^2 : b = 0\}$

    c) $V = \mathcal{M}_{2\times2}$ y $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a = d, \text{ y } b = c \right\}$

    d) $V =$ el plano de R^3 que pasa por el origen y tiene como vector normal a $\mathbf{n} = (-1 \;\; 2 \;\; 5)^T$ y $W = \{\mathbf{v} \in R^2 : \mathbf{v} = \lambda(-2 \;\; 5)^T\}$

    e) $V = \{(x \;\; y \;\; z \;\; w)^T : x - y + 2z = 0\}$ y $W = \{(x \;\; y \;\; z \;\; w)^T : x + 2z - w = 0\}$

22. Complete las demostraciones de los Teoremas 16 y 17.

23. Para cada uno de los siguientes casos, encuentre, si existe, $S + T$, $T - S$, $-2T$, T^{-1}, $S \circ T$ y $T \circ S$ y sus matrices asociadas respecto a las bases que usted elija.

    a) $T : R^3 \longrightarrow R^2$ tal que $T(a \;\; b \;\; c)^T = (2a \;\; b-3c)^T$; $S : R^3 \longrightarrow R^2$ tal que $S(a \;\; b \;\; c)^T = (a \;\; 2b+c)^T$.

    b) $T : R^3 \to R^3$ tal que $T(a \;\; b \;\; c)^T = (a \;\; 2b+c \;\; -c)^T$; $S : R^3 \to R^3$ tal que $S(a \;\; b \;\; c)^T = (-a \;\; 2b \;\; 3c)^T$.

    c) $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_1$ tal que $T(p(x)) = p'(x)$; $S : \mathcal{P}_1 \longrightarrow \mathcal{P}_2$ tal que $S(p(x)) = xp(x)$.

    d) $T : \mathcal{M}_{2\times2} \to \mathcal{P}_2$ tal que $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a+bx+cx^2$; $S : \mathcal{P}_2 \to \mathcal{P}_2$ tal que $S(a+bx+cx^2) = a-cx^2$.

    e) $T : R^3 \longrightarrow R^2$ tal que $T(a \;\; b \;\; c)^T = (a \;\; 2b-3c)^T$; $S : R^2 \longrightarrow R^3$ tal que $S(a \;\; b)^T = (-3a \;\; b+a \;\; b)^T$

    f) $T : R^3 \longrightarrow \mathcal{P}_2$ tal que $T(a \;\; b \;\; c)^T = a+2bx-3cx^2$; $S : \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ tal que $S(a+bx+cx^2) = a-cx^2$

24. Demuestre el Teorema 18. ***(AYUDA:*** *Ver demostración del Teorema 5 del Capítulo 3.)*

25. Si $T : V \to W$ es un isomorfismo y definimos $S : W \to V$, de tal forma que $S(\mathbf{w}) = \mathbf{v}$, cuando $T(\mathbf{v}) = \mathbf{w}$, demuestre que S es la inversa de T.

26. Demuestre el Teorema 20. ***(AYUDA:*** *Para la primera parte, use los resultados de los Teorema 12 y 19; para la segunda, use una técnica similar a la demostración del Numeral 2 del Teorema 16 del Capítulo 4.)*

27. Conteste falso o verdadero a cada una de las siguientes afirmaciones, justificando su respuesta.

    a) Toda transformación lineal es una función.
    b) Si $T(\mathbf{0}) = \mathbf{0}$, T es una transformación lineal.
    c) Si $T : V \longrightarrow W$ es una transformación lineal y $\dim V = \dim W$, entonces T es invertible.
    d) Si $T : V \longrightarrow W$ es una transformación lineal y T es invertible, entonces $\dim V = \dim W$.
    e) Si $T : V \longrightarrow W$ es una transformación lineal inyectiva y $\dim V = \dim W$, entonces T es invertible.
    f) Si $T : V \longrightarrow W$ es una transformación lineal sobreyectiva y $\dim V = \dim W$, entonces T es inyectiva.
    g) Si $T : V \longrightarrow W$ es una transformación lineal y $\dim V < \dim W$, entonces T no es inyectiva
    h) Si $T : V \longrightarrow W$ es una transformación lineal y $\dim V < \dim W$, entonces T no es sobreyectiva.
    i) Si $T : V \longrightarrow W$ es una transformación lineal y T es inyectiva, entonces $\dim V \leq \dim W$.
    j) Si $T : V \longrightarrow W$ es una transformación lineal y T es sobreyectiva, entonces $\dim V < \dim W$.
    k) Si $T : V \longrightarrow W$ es una transformación lineal y $\mathbf{v} \in Nu(T)$, entonces $\mathbf{v} \in N_{A_T}$.
    l) Si $T : V \longrightarrow W$ es una transformación lineal y $\mathbf{w} \in Im(T)$, entonces, para alguna base $\mathcal{B}$ de W, $[\mathbf{w}]_{\mathcal{B}} \in N_{A_T}$.
    m) Si $T : V \longrightarrow W$ es una transformación lineal, $S : V \longrightarrow Im(T)$ tal que $S(\mathbf{v}) = T(\mathbf{v})$ es una transformación lineal invertible.
    n) Si $T : V \longrightarrow W$ es una transformación lineal, $S : Nu(T) \longrightarrow W$ tal que $S(\mathbf{v}) = T(\mathbf{v})$ es la transformación lineal nula.
    ñ) Si $T : \mathcal{M}_{2\times 2} \longrightarrow \mathcal{P}_3$ es una transformación lineal sobreyectiva , entonces T es inyectiva.
    o) La proyección ortogonal en un subespacio es una transformación lineal inyectiva del espacio vectorial en el subespacio.
    p) La proyección ortogonal en un subespacio es una transformación lineal sobreyectiva del espacio vectorial en el subespacio.
    q) La rotación de 60° en el sentido contrario a la manecilla del reloj es un isomorfismo de R^2 en R^2.
    r) La translación de vectores en el plano, $T : R^2 \longrightarrow R^2$ tal que $T(\mathbf{x}) = \mathbf{x} + \mathbf{a}$, $\mathbf{a} \in R^2$ fijo, no es una transformación lineal.
    s) Si $\mathbf{u} = \mathbf{v}$ implica que $T(\mathbf{u}) = T(\mathbf{v})$, entonces T es una transformación lineal inyectiva.
    t) Siempre que se puedan calcular las composiciones $S \circ T$ y $T \circ S$, se tiene que una es la inversa de la otra.
    u) $(T \circ S)(\mathbf{v}) = A_T A_S \mathbf{v}$, si A_T y A_S, las matrices asociadas a las transformaciones lineales, son las matrices asociadas respecto a las bases canónicas de los dominios y codominios.
    v) La matriz que define una transformación matricial es la matriz asociada a la transformación lineal en cualquier base.
    w) Una transformación lineal puede tener infinitas matrices asociadas.
    x) Dos matrices semejantes representan la misma transformación lineal en diferentes bases.
    y) Dados dos espacios vectoriales V y W, $\mathcal{L}(V, W)$, el conjunto de las transformaciones lineales de V en W es un espacio vectorial.
    z) Una función lineal (f(x)=mx+b) es una transformación lineal.

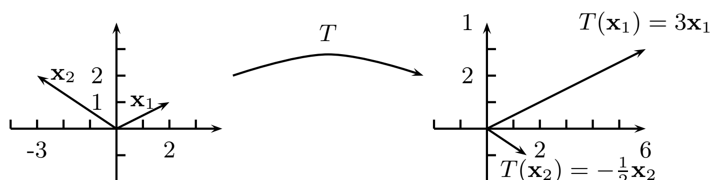
Capítulo 6

VALORES Y VECTORES PROPIOS

6.1. Introducción

Aunque, en general, la imagen de un vector bajo una transformación de un espacio vectorial en si mismo no es un vector paralelo al vector inicial, existen vectores especiales para los cuales la acción de la transformación es muy sencilla: el vector y su imagen tienen la misma dirección. Estos vectores especiales los denominamos vectores propios y conocerlos son de gran ayuda en el análisis de la transformación puesto que en la dirección de ellos, la transformación sólo "encoge" o "estira" los vectores, cambiando, en algunos casos, el sentido de ellos.

Por ejemplo, la transformación matricial $T : R^2 \longrightarrow R^2$, definida por $T(\mathbf{x}) = A\mathbf{x}$, donde $A = \begin{pmatrix} 1{,}5 & 3 \\ 1 & 1 \end{pmatrix}$, en la dirección de $\mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, "estira" los vectores, puesto que $T(\mathbf{x}_1) = A\mathbf{x}_1 = 3\mathbf{x}_1$ y, en la dirección de $\mathbf{x}_2 = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$, "encoge" los vectores y los cambia de sentido, puesto que $T(\mathbf{x}_2) = A\mathbf{x}_2 = -\frac{1}{2}\mathbf{x}_2$, lo cual ilustramos en el siguiente gráfico.



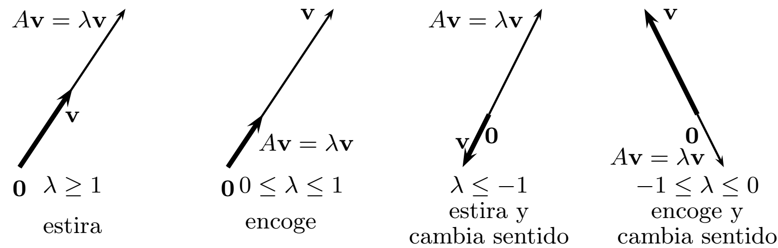
Esta característica de los vectores propios es muy importante en el estudio de ecuaciones diferenciales en general, y en el análisis de los sistemas dinámicos, tanto discretos como continuos, en particular. En la práctica, los vectores propios suministran información crítica en el diseño de ingeniería y en diversas aplicaciones de la física y la química.

Por la relación entre las transformaciones y las matrices que estudiamos en el capítulo anterior, estudiaremos los valores y vectores propios de una transformación como una propiedad de las matrices asociadas. Así por ejemplo, si $\mathbf{v}$ es un vector tal que $\mathbf{v} = A\mathbf{v}$, decimos que $\mathbf{v}$ es un *vector invariante o estacionario* para la trasformación definida por la matriz A. Si el vector $\mathbf{u}$ tiene la propiedad de ser paralelo a su imagen mediante la transformación definida por la matriz A, digamos que $\mathbf{u} = 2A\mathbf{u}$, decimos que $\mathbf{u}$ es un vector propio de la matriz A o equivalentemente, de la transformación definida por A.

6.2. Conceptos Básicos

Como mencionamos en la introducción, aunque los vectores propios son vectores paralelos a sus imágenes bajo una transformación, por la relación que existe entre las transformaciones lineales y las matrices, definimos primero este y otros conceptos relacionados, respecto de una matriz para más adelante hacerlo respecto de una transformación. Es de anotar que, como estos conceptos tienen sentido sólo para transformaciones que van de un espacio vectorial en si mismos, en este capítulo, estaremos hablando sólo de *matrices cuadradas.*

Definición 1 [*Vector y valor propio de una matriz*]. Dada A una matriz $n \times n$, decimos que $\lambda \in R$ es un *valor propio*[1] de A y que $\mathbf{v} \in R^n$, $\mathbf{v} \neq \mathbf{0}$, es un *vector propio* de A asociado a λ, si y solo si, $A\mathbf{v} = \lambda\mathbf{v}$. [2]



Observemos que la exclusión del vector cero como vector propio es natural, ya que si, en la ecuación vectorial $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} = \mathbf{0}$, entonces λ puede ser cualquier número real.

Observemos también que a un valor propio le corresponden infinitos vectores propios. En efecto, si $\mathbf{v}$ es un vector propio asociado al valor propio λ, el vector $\mathbf{u} = \mu\mathbf{v}$, para cualquier $\mu \in R - \{0\}$, también es un vector propio asociado a λ ya que $A\mathbf{u} = A(\mu\mathbf{v}) = \mu(A\mathbf{v}) = \mu(\lambda\mathbf{v}) = \lambda(\mu\mathbf{v}) = \lambda\mathbf{u}$.

Ejemplo 1. Consideremos $A = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ -1 & 1 & -2 \end{pmatrix}$. Verifiquemos que $\mathbf{u} = \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$ son vectores propios de A asociados a los valores propios $\lambda_1 = 0$ y $\lambda_2 = -2$, respectivamente.

$$A\mathbf{u} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ -1 & 1 & -2 \end{pmatrix}\begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0\begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} = \lambda_1\mathbf{u}$$

$$A\mathbf{v} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ -1 & 1 & -2 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix} = -2\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \lambda_2\mathbf{v}$$

□

De la definición anterior, tenemos que un vector propio asociado a λ es una solución no nula de $A\mathbf{x} = \lambda\mathbf{x}$ o lo que es equivalente, un vector propio, asociado a λ, es una solución **no trivial** del sistema de ecuaciones lineales homogéneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

Por el Corolario 23.1 del Capítulo 3, el anterior sistema homogéneo tiene soluciones diferentes a la trivial, si y solo si, $\det(A - \lambda I) = 0$. Así que, los valores propios de A son los valores reales de λ que hacen que $\det(A - \lambda I) = 0$ y los vectores propios asociados a λ son las soluciones no nulas de $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Este razonamiento es la demostración del siguiente teorema, el cual permite caracterizar separadamente los valores propios (parte no lineal) y los vectores propios asociados a un valor propio (parte lineal).

[1]Es importante anotar que los valores propios son números reales porque estamos trabajando en espacios vectoriales **reales**, pero podrían ser complejos cuando consideremos espacios vectoriales donde los escalares son los números complejos.

[2]Los valores y vectores propios también se conocen como *valores y vectores característicos* o *eigenvalores y eigenvectores.*

Teorema 1 [*Caracterización de los vectores y valores propios de una matriz*].
Sea A una matriz cuadrada. Entonces

1. Un escalar λ es valor propio de A, si y solo si, $\det(A - \lambda I) = 0$ y $\lambda \in R$.
2. Un vector $\mathbf{v}$ es un vector propio de A asociado al valor propio λ, si y solo si, $\mathbf{v}$ es una solución no trivial de $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

Ejemplo 2. Encontremos todos los valores y vectores propios de la matriz A del Ejemplo 1.

Aunque en general no es fácil el cálculo de un determinante, podemos ver que $\det(A - \lambda I) = -\lambda^3 - 2\lambda^2$. Así que si $\det(A - \lambda I) = 0$, entonces $\lambda = 0$ ó $\lambda = -2$, de donde concluimos que $\lambda_1 = 0$ y $\lambda_2 = -2$ son los valores propios de A, lo cual ya habíamos verificado en el Ejemplo 1.

Al resolver $(A - 0I)\mathbf{x} = A\mathbf{x} = \mathbf{0}$, tenemos que el conjunto solución es $S_1 = \left\{ z \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}, \ z \in R \right\}$. Por tanto, los vectores propios de A asociados al valor propio $\lambda_1 = 0$ son todos los múltiplos no nulos de $\begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}$.

Y al resolver $(A + 2I)\mathbf{x} = \mathbf{0}$, tenemos que el conjunto solución es $S_2 = \left\{ z \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \ z \in R \right\}$. Por tanto, los vectores propios de A asociados al valor propio $\lambda_2 = -2$ son todos los múltiplos no nulos de $\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$. □

Notemos que el vector $\mathbf{0}$, por definición, no es vector propio, mientras que el escalar 0 si puede ser valor propio. En particular, tenemos que si la matriz A tiene un valor propio igual a 0 es porque $\det(A - 0I) = \det A = 0$, lo que es equivalente a que la matriz sea no invertible, resultado que consignamos en el siguiente teorema.

Teorema 2 [*Valores propios e invertibilidad*].
Una matriz A es invertible, si y solo si, $\lambda = 0$ **no** es un valor propio de A.

Como mencionamos antes, los valores propios de A son las soluciones reales de la ecuación $\det(A - \lambda I) = 0$. Además, se puede demostrar que si A es una matriz $n \times n$, entonces $\det(A - \lambda I)$ es un polinomio de grado n, cuyo coeficiente principal[3] es $(-1)^n$. Por la importancia que tiene este polinomio, le damos un nombre especial.

Definición 2 [*Polinomio característico*]. Dada una matriz A, decimos que $p(\lambda) = \det(A - \lambda I)$ es el *polinomio característico de A*.

Ejemplo 3. Calculemos el polinomio característico de la matriz $A = \begin{pmatrix} 3 & -1 & 7 \\ 0 & 5 & 2 \\ 0 & 0 & -1 \end{pmatrix}$.

Aunque en general este polinomio no es fácil de calcular, en este caso, por ser A una matriz triangular, tenemos que

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 3-\lambda & -1 & 7 \\ 0 & 5-\lambda & 2 \\ 0 & 0 & -1-\lambda \end{vmatrix} = (3-\lambda)(5-\lambda)(-1-\lambda) = -\lambda^3 + 7\lambda^2 - 7\lambda - 15.$$

□

Ejemplo 4. La matriz $A = \begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$ no tiene valores propios, ya que su polinomio característico, $p(\lambda) = \lambda^2 + 1$, no tiene raíces reales. □

[3]Se llama coeficiente principal de un polinomio al coeficiente del término de mayor potencia en el polinomio

Recordemos que el Teorema Fundamental del Álgebra establece que todo polinomio de coeficientes reales de grado n tiene n raíces complejas, contando multiplicidades[4]. Por tanto, si A es una matriz $n \times n$, podemos concluir que A tiene a lo más n valores propios, contando multiplicidades.

Definición 3 [*Multiplicidad algebraica de un valor propio*]. A la multiplicidad de una raíz real λ del polinomio característico de la matriz A le llamamos *multiplicidad algebraica* del valor propio λ.

Ejemplo 5. Las multiplicidades algebraicas de los valores propios de la matriz del Ejemplo 3 son todas 1. □

Ejemplo 6. Si $A = \begin{pmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2 \end{pmatrix}$, después de una serie de cálculos, tenemos que su polinomio característico es $p(\lambda) = -\lambda^3 + 5\lambda^2 - 8\lambda + 4 = (1-\lambda)(\lambda-2)^2$, de donde concluimos que $\lambda_1 = 1$ es un valor propio de multiplicidad algebraica 1 y $\lambda_2 = 2$ es un valor propio de multiplicidad algebraica 2. □

Cuáles son las multiplicidades algebraicas de los valores propios de la matriz del Ejemplo 2?

Como lo mencionamos antes, podemos ver que los vectores propios asociados a un valor propio λ son las soluciones no triviales del sistema homogéneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$. En otras palabras, el conjunto de los vectores propios asociados a un valor propio λ, junto con el vector $\mathbf{0}$, coincide con el espacio nulo de la matriz $(A - \lambda I)$ y le daremos un nombre especial.

Definición 4 [*Espacio propio*]. Dada A una matriz de tamaño $n \times n$ y $\lambda \in R$ un valor propio de A, definimos E_λ, el *espacio propio de A asociado a λ*, como el conjunto conformado por todos los vectores propios de A asociados a λ y el vector $\mathbf{0}$ de R^n; es decir,

$$E_\lambda = N_{(A-\lambda I)}.$$

Ejemplo 7: Debido a que los valores propios de la matriz A del Ejemplo 2 son $\lambda_1 = 0$ y $\lambda_2 = 2$, los espacios propios E_0 y E_{-2} son los conjuntos solución de $A\mathbf{x} = \mathbf{0}$ y $(A + 2I)\mathbf{x} = \mathbf{0}$, encontrados en el Ejemplo 2. En otras palabras, los espacios propios de A son $E_0 = \left\{ z \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}, z \in R \right\}$ y $E_{-2} = \left\{ z \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, z \in R \right\}$.□

Por su definición, los espacios propios de una matriz $n \times n$ asociados a sus valores propios son subespacios vectoriales de R^n. A las dimensiones de estos subespacios, les llamamos de una manera particular, ya que nos permiten conocer otros detalles de la matriz.

Definición 5 [*Multiplicidad geométrica de un valor propio*]. Dada una matriz A de tamaño $n \times n$ y λ un valor propio de A, definimos como *multiplicidad geométrica del valor propio* λ a la dimensión del espacio propio E_λ.

Ejemplo 8. Dada la matriz $A = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 1 & 2 \\ -1 & -2 & 0 \end{pmatrix}$, determinemos las multiplicidades geométricas de sus valores propios; es decir, las dimensiones de sus espacios propios.

Después de una serie de operaciones, tenemos que el polinomio característico de A es

$$p(\lambda) = -\lambda^3 + 5\lambda^2 - 7\lambda + 3 = (\lambda - 1)^2(3 - \lambda),$$

así que sus valores propios son $\lambda_1 = 1$ de multiplicidad algebraica 2 y $\lambda_2 = 3$ de multiplicidad algebraica 1.

Al encontrar los conjuntos solución de $(A - I)\mathbf{x} = \mathbf{0}$ y $(A - 3I)\mathbf{x} = \mathbf{0}$, tenemos que los espacios propios son

$$E_1 = \left\{ x \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, x \in R \right\} \quad \text{y} \quad E_3 = \left\{ z \begin{pmatrix} -5 \\ -2 \\ 3 \end{pmatrix}, z \in R \right\},$$

[4] Se dice que a es una raíz del polinomio $p(x)$ con multiplicidad r, si y solo si, $p(x) = (x-a)^r q(x)$ y $q(a) \neq 0$.

de donde podemos concluir que las multiplicidades geométricas de los espacios propios de A son ambas 1.□

Por las definiciones de multiplicidad algebraica y geométrica, es claro que ellas son mayores o iguales a uno. Además, existe una relación entre las multiplicidades algebraica y geométrica de un mismo valor propio que aunque no es difícil de demostrar, no está al alcance del nivel de este texto, pero por su importancia la expresamos en el siguiente teorema.

Teorema 3 [*Relación entre las multiplicidades de un valor propio*].
Si λ es un valor propio de una matriz $n \times n$, A, entonces

$$1 \leq \textit{multiplicidad geométrica de } \lambda \leq \textit{multiplicidad algebraica de } \lambda \leq n\,.$$

Del Teorema 3, es fácil concluir que si la multiplicidad algebraica de un valor propio es 1, las dos multiplicidades son iguales.

Ejemplo 9. Observemos que las multiplicidades algebraica y geométrica del valor propio $\lambda_1 = 1$ de la matriz A del Ejemplo 8 son ambas 1, mientras que las multiplicidades algebraica y geométrica del valor propio $\lambda_2 = 3$ son 2 y 1, respectivamente, satisfaciendo el Teorema 3. □

6.3. Cálculo de Valores y Vectores Propios

El Teorema 1 y los comentarios anteriores nos insinúan un procedimiento para encontrar los vectores y valores propios de una matriz, el cual tiene sentido práctico para valores pequeños de n, el orden de la matriz.

Procedimiento para calcular valores y vectores propios de una matriz A.

1. Calcular el polinomio característico $p(\lambda) = \det(A - \lambda I)$.
2. Calcular las raíces reales de $p(\lambda)$ (*valores propios de* A).
3. Resolver el sistema homogéneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$, para cada valor de λ (*vectores propios de* A *asociados a* λ).

Ejemplo 10. Dada la matriz A, encontremos los valores y vectores propios de $A = \begin{pmatrix} 3 & -2 \\ -3 & 2 \end{pmatrix}$.

El polinomio característico de A es

$$p(\lambda) = \det \begin{pmatrix} 3-\lambda & -2 \\ -3 & 2-\lambda \end{pmatrix} = \lambda^2 - 5\lambda.$$

Como las raíces de este polinomio son $\lambda = 0$ y $\lambda = 5$, los valores propios de A son $\lambda_1 = 0$ y $\lambda_2 = 5$, ambos con multiplicidad algebraica 1.

Al resolver el sistema

$$(A - 0I)\mathbf{x} = A\mathbf{x} = \begin{pmatrix} 3 & -2 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} ,$$

tenemos que los vectores propios de A asociados a $\lambda = 0$ son los vectores no nulos $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ tales que $3x_1 - 2x_2 = 0$; es decir,

$$E_0 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : 3x_1 - 2x_2 = 0 \right\} = \left\{ \begin{pmatrix} x_1 \\ \frac{3}{2}x_1 \end{pmatrix} : x_1 \in R \right\} = Gen \left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}.$$

Y al resolver el sistema

$$(A - 5I)\mathbf{x} = \begin{pmatrix} -2 & -2 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

tenemos que los vectores propios de A asociados a $\lambda = 5$ son los vectores no nulos $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ tales que $-2x_1 - 2x_2 = 0$; es decir,

$$E_5 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 + x_2 = 0 \right\} = \left\{ \begin{pmatrix} x_1 \\ -x_1 \end{pmatrix} : x_1 \in R \right\} = Gen \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

□

Ejemplo 11. Sea $A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. Después de hacer varios cálculos para calcular y factorizar el polinomio característico de A, obtenemos que

$$p(\lambda) = \det(A - \lambda I) = (2 - \lambda)(\lambda^2 + 1) \ ,$$

cuyas raíces son $\lambda = 2$ y $\lambda = \pm i$, de donde concluimos que la matriz A sólo tiene un valor propio, $\lambda = 2$, que tiene multiplicidad algebraica (y por lo tanto, la geométrica) igual a 1.

Al resolver el sistema

$$(A - 2I)\mathbf{x} = \begin{pmatrix} -2 & 1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

tenemos que los vectores propios de A asociados a $\lambda = 2$ son los vectores no nulos $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ tales que $-2x_1 + 2x_2 = 0$ y $-x_1 - 2x_2 = 0$; es decir,

$$E_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 = x_2 = 0 \right\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} : x_3 \in R \right\} = Gen \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

lo que corrobora que la multiplicidad geométrica también es 1. □

Es importante hacer notar que, en general, encontrar el polinomio característico de una matriz no es fácil (recordemos que dicho polinomio resulta del cálculo de un determinante) y que encontrar las raíces de un polinomio también es un problema difícil de resolver. Así, el procedimiento sugerido para encontrar los valores y vectores propios de una matriz puede resultar muy difícil de usar[5], aunque sí es un procedimiento sencillo para **verificar** que un número es o no un valor propio y/o que un vector dado es o no un vector propio.

Por el Teorema 13 del Capítulo 3, sabemos que si la matriz es triangular, su determinante es simplemente el producto de los elementos de la diagonal. Así que calcular el polinomio característico y los valores propios de una matriz triangular es fácil (ver Ejemplo 3).

Teorema 4 [*Valores propios de una matriz triangular*].
Los valores propios de una matriz triangular son los elementos de la diagonal de la matriz.

Demostración: Si $A = (a_{ij})$ es una matriz triangular, su polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda).$$

Así que los valores propios de A, que son las soluciones de $p(\lambda) = 0$ son $a_{11}, a_{22}, \cdots a_{nn}$, los cuales son los elementos de la diagonal de A. □

[5]En la práctica, para calcular los valores propios de una matriz, se utilizan métodos numéricos como el métodos de las potencias, el método de Raleigh y el método QR de Francis (ver MARTINEZ H.J y PEREZ R. Introducción al Álgebra Lineal Numérica. Ed. Universidad del Cauca, 1990).

6.4. Independencia de los vectores propios

Las preguntas naturales sobre los vectores propios son si forman un conjunto de vectores linealmente independientes y si constituyen un conjunto generador de R^n, cuyas respuestas parciales se encuentran en el siguiente teorema.

Teorema 5 [*Independencia lineal y bases de R^n de vectores propios*].
Si $\lambda_1, \lambda_2, \ldots, \lambda_k$ son valores propios distintos de una matriz A de tamaño $n \times n$, entonces

1. El conjunto de vectores propios $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ de A asociados a los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_k$, respectivamente, es *l.i.*

2. Si $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ son bases de los espacios propios $E_{\lambda_1}, E_{\lambda_2}, \ldots, E_{\lambda_k}$, respectivamente, entonces

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_k$$

   es un conjunto *l.i.*

3. Si $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ son bases de los espacios propios $E_{\lambda_1}, E_{\lambda_k}, \ldots, E_{\lambda_k}$, respectivamente, y

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_k$$

   tiene n elementos, $\mathcal{B}$ es una base de R^n.

Demostración: Demostremos el primer resultado, usando inducción matemática finita sobre k, el número de valores propios distintos de la matriz A ($k \leq n$), y dejamos la demostración de los otros resultados como ejercicio para el lector.

El resultado es trivialmente válido para $k = 1$. Demostremos el resultado para $k = 2$. Supongamos que

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 = \mathbf{0}. \tag{6.1}$$

Multiplicando ambos lados por la matriz A, tenemos

$$\begin{aligned} \alpha_1 A\mathbf{v}_1 + \alpha_2 A\mathbf{v}_2 &= \mathbf{0} && (6.2) \\ \alpha_1(\lambda_1 \mathbf{v}_1) + \alpha_2(\lambda_2 \mathbf{v}_2) &= \mathbf{0}. && (6.3) \end{aligned}$$

Multiplicando (6.1) por λ_2 y restando el resultado de (6.3), obtenemos

$$\alpha_1(\lambda_1 - \lambda_2)\mathbf{v}_1 = \mathbf{0}.$$

Por hipótesis $\lambda_1 \neq \lambda_2$ y puesto que $\mathbf{v}_1$ es un vector propio, $\mathbf{v}_1 \neq 0$, entonces $\alpha_1 = 0$. Reemplazando este valor en (6.1), como $\mathbf{v}_2 \neq 0$ por ser un vector propio, concluimos que $\alpha_2 = 0$; por lo tanto, $\mathbf{v}_1$ y $\mathbf{v}_2$ son *l.i.*

Supongamos que el resultado es válido para $k = r < n$; es decir, $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r\}$ es un conjunto de vectores *l.i.* y demostremos que el resultado es válido para $k = r + 1 \leq n$. Supongamos que

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_{r+1} \mathbf{v}_{r+1} = \mathbf{0}. \tag{6.4}$$

Multiplicando ambos lados por la matriz A, tenemos

$$\begin{aligned} \alpha_1 A\mathbf{v}_1 + \alpha_2 A\mathbf{v}_2 + \ldots + \alpha_{r+1} A\mathbf{v}_{r+1} &= \mathbf{0} && (6.5) \\ \alpha_1(\lambda_1 \mathbf{v}_1) + \alpha_2(\lambda_2 \mathbf{v}_2) + \cdots + \alpha_{r+1}(\lambda_{r+1} \mathbf{v}_{r+1}) &= \mathbf{0} && (6.6) \end{aligned}$$

Multiplicando (6.4) por λ_{r+1} y restando el resultado de (6.6), obtenemos

$$\alpha_1(\lambda_1 - \lambda_{r+1})\mathbf{v}_1 + \alpha_2(\lambda_2 - \lambda_{r+1})\mathbf{v}_2 + \cdots + \alpha_r(\lambda_r - \lambda_{r+1})\mathbf{v}_r = \mathbf{0}.$$

Como por hipótesis $\lambda_1, \lambda_2, \ldots, \lambda_r \neq \lambda_{r+1}$ y por hipótesis de inducción el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ es *l.i.*, entonces $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$. Reemplazando estos valores en (6.4), tenemos que por ser $\mathbf{v}_{r+1}$ un vector propio, $\mathbf{v}_{r+1} \neq 0$, de donde concluimos que $\alpha_{r+1} = 0$; por lo tanto, el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{r+1}\}$ es *l.i.* $\square$

Ejemplo 12. Si $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 \end{pmatrix}$, entonces, después de muchos cálculos para determinar el polinomio característico de A y factorizarlo, tenemos que

$$p(\lambda) = \lambda^4 - 5\lambda^3 + 9\lambda^2 - 7\lambda + 2 = (\lambda - 2)(\lambda - 1)^3$$

y por lo tanto, sus valores propios son $\lambda_1 = 2$ de multiplicidad algebraica 1 y $\lambda_2 = 1$ de multiplicidad algebraica 3.

Al resolver los sistemas homogéneos $(A - 2I)\mathbf{x} = \mathbf{0}$ y $(A - I)\mathbf{x} = \mathbf{0}$, encontramos que los espacios propios de A son $E_2 = Gen\left\{\begin{pmatrix} 4 \\ 1 \\ 0 \\ 3 \end{pmatrix}\right\}$ y $E_1 = Gen\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\}$. Así que la multiplicidad geométrica de los dos valores propios es 1. Además, es fácil concluir que los dos vectores de los conjuntos generadores de los espacios propios E_2 y E_1 forman un conjunto *l.i.*, ya que uno no es múltiplo del otro; por lo tanto, cualquier conjunto formado por un vector propio asociado a $\lambda_1 = 2$ y un vector propio asociado a $\lambda_2 = 1$ es *l.i.* $\square$

Observemos que si todas las raíces del polinomio característico de una matriz son reales y cada una de ellas tiene asociado tantos vectores propios *l.i.* como su multiplicidad algebraica, se satisface la hipótesis del tercer resultado del teorema anterior. Esta observación es la base de la demostración del siguiente corolario, la cual dejamos como ejercicio para el lector.

Corolario 5.1 [*Condición suficiente para una base de R^n formada por vectores propios*].
Si cada uno de los valores propios de una matriz A de tamaño $n \times n$ tiene su multiplicidad algebraica y geométrica iguales y la suma de sus multiplicidades algebraicas es n, existe una base de R^n formada por vectores propios de la matriz A.

Ejemplo 13. Si $A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$, entonces, después de una serie de cálculos, obtenemos que el polinomio característico de A es $p(\lambda) = -\lambda^3 + 12\lambda^2 - 21\lambda + 10 = -(\lambda - 10)(\lambda - 1)^2$ y por lo tanto, sus valores propios son $\lambda_1 = 10$ de multiplicidad algebraica 1 y $\lambda_2 = 1$ de multiplicidad algebraica 2.

Al resolver los sistemas homogéneos $(A - 10I)\mathbf{x} = \mathbf{0}$ y $(A - I)\mathbf{x} = \mathbf{0}$, encontramos que los espacios propios de A son $E_{10} = Gen\left\{\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}\right\}$ y $E_1 = Gen\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}\right\}$, de donde concluimos que los valores propios λ_1 y λ_2 tienen multiplicidad geométrica 1 y 2, respectivamente. Como las multiplicidades algebraicas y geométricas de los valores propios de A son iguales, por el Corolario 5.1, concluimos que existe una base de R^3 formada por vectores propios de A.

Además, por el Resultado 3 del Teorema 5, podemos concluir que

$$\mathcal{B} = \left\{\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}\right\}$$

es una base de R^3 formada por vectores propios de A. $\square$

Por el comentario que hicimos sobre el Teorema 3 y recordando que, por el Teorema Fundamental del Álgebra, si un polinomio de grado n tiene n raíces distintas, entonces la multiplicidad de cada una de ellas debe ser 1,

un caso particular del resultado anterior, por su importancia práctica, es el que enunciamos en el siguiente corolario.

Corolario 5.2 [*Otra condición suficiente para una base de R^n formada por vectores propios*].
Si una matriz A de tamaño $n \times n$ tiene n valores propios distintos, existe una base de R^n formada por vectores propios de la matriz A.

Ejemplo 14. Determine si existe una base de R^3 formada por vectores propios de la matriz

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 3 & 2 & 0 \\ 1 & 3 & 5 \end{pmatrix}.$$

Como A es una matriz triangular, por el Teorema 4, sus valores propios son $\lambda_1 = -1$, $\lambda_2 = 2$ y $\lambda_3 = 5$, todos de multiplicidad algebraica 1. Por el Corolario 5.2, concluimos que si existe una base de R^3 formada por vectores propios de A. En efecto, al resolver los sistemas homogéneos $(A+I)\mathbf{x} = \mathbf{0}$, $(A-2I)\mathbf{x} = \mathbf{0}$ y $(A-5I)\mathbf{x} = \mathbf{0}$, encontramos que los espacios propios de A son $E_{-1} = Gen\left\{\begin{pmatrix} -3 \\ 3 \\ -1 \end{pmatrix}\right\}$, $E_2 = Gen\left\{\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right\}$ y $E_5 = Gen\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$. Por el Resultado 3 del Teorema 5, concluimos que $\mathcal{B} = \left\{\begin{pmatrix} -3 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ es una base de R^3 formada por vectores propios de A. □

Para terminar esta sección de conceptos básicos, veamos la relación que existe entre los valores y vectores propios de una matriz y los de la matriz que resulta al aplicarle a la matriz original algunas de las operaciones de matrices estudiadas en el Capítulo 3.

Teorema 6 [*Propiedades de los valores y vectores propios*].
Sea A una matriz de orden $n \times n$ y $\mathbf{v}$ un vector propio de A asociado al valor propio λ. Entonces,

1. λ es un valor propio de A^T.

2. Si A es invertible, entonces $\lambda \neq 0$ y $\mathbf{v}$ es un vector propio de A^{-1} asociado al valor propio $1/\lambda$.

3. $\mathbf{v}$ es un vector propio de A^k asociado al valor propio λ^k.

Demostración: La demostración de estos resultados está basada en las propiedades de los determinantes y la caracterización de los valores propios como raíces del polinomio característico. Demostraremos la segunda propiedad y las otras las dejamos como ejercicio para el lector.

2. Si A es invertible, $\lambda = 0$ no puede ser un valor propio de A, porque si lo fuere, $\det(A - 0I) = \det(A) = 0$ y, por el Corolario 15.1 del Capítulo 3, la matriz A no sería invertible. Además, si $\mathbf{v}$ es un vector propio de A, asociado a $\lambda \neq 0$, entonces $A\mathbf{v} = \lambda\mathbf{v}$ o equivalentemente, $\frac{1}{\lambda}\mathbf{v} = A^{-1}\mathbf{v}$; es decir, $\mathbf{v}$ es un vector propio de A^{-1} asociado al valor propio $1/\lambda$.

Ejemplo 15. Verifiquemos los resultados del Teorema 6 para la matriz $A = \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix}$.

El polinomio característico de A y A^T es $p(\lambda) = \lambda^2 - 9$ y por tanto, los valores propios son $\lambda_1 = 3$, $\lambda_2 = -3$, para ambas matrices.

Por ser A una matriz 2×2, no es difícil calcular su inversa, obteniendo que $A^{-1} = \begin{pmatrix} 1/9 & 2/9 \\ 4/9 & -1/9 \end{pmatrix}$, así que su polinomio característico es $p(\lambda) = \lambda^2 - \dfrac{1}{9}$ y por tanto, sus valores propios son $\lambda_1 = 1/3$, $\lambda_2 = -1/3$, los inversos de los valores propios de A.

Al resolver los sistemas homogéneos $(A-3I)\mathbf{x}=\mathbf{0}$ y $(A+3I)\mathbf{x}=\mathbf{0}$, encontramos que los espacios propios de A son $E_3 = Gen\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ y $E_{-3} = Gen\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\}$. Verifiquemos que tanto $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, como $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ son vectores propios de A^{-1}, correspondientes a los inversos de los valores propios de A, lo que, además, nos permite concluir que todos los vectores propios de A son vectores propios de A^{-1}:

$$\begin{aligned}
\begin{pmatrix} 1/9 & 2/9 \\ 4/9 & -1/9 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 1/9 & 2/9 \\ 4/9 & -1/9 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} &= \begin{pmatrix} -1/3 \\ 2/3 \end{pmatrix} = \frac{-1}{3}\begin{pmatrix} 1 \\ -2 \end{pmatrix}.
\end{aligned}$$

Finalmente, para ilustrar él último resultado del Teorema 6, veamos que

$$A^3\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 9 & 18 \\ 36 & -9 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 27 \\ 27 \end{pmatrix} = 3^3\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Es decir, que $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ es también vector propio de A^3, pero asociado al valor propio $3^3 = 27$. □

En el ejemplo anterior, con el propósito de ilustrar los resultados del Teorema 6, calculamos A^{-1} y A^3 lo cual es innecesario en la práctica. La importancia del Teorema 6 radica precisamente en poder calcular los valores y vectores propios de A^{-1} y A^3, conociendo los de A, sin tener que calcular estas matrices.

Otro resultado importante similar a los anteriores es la caracterización de la invertibilidad de una matriz dada por sus valores propios, la cual explicitamos en el siguiente corolario, cuya demostración dejamos como ejercicio para el lector.

Corolario 6.1 [*Caracterización de las matrices **no** invertibles*].
Una matriz A de tamaño $n \times n$ es **no** invertible, si y solo si, $\lambda = 0$ es un valor propio de A.

Observemos que otra forma de expresar el Corolario 6.1 es diciendo que una matriz cuadrada es invertible, si y solo si, todos sus valores propios son diferentes de cero, o equivalentemente, si y solo si, $\lambda = 0$ no es un valor propio de A.

6.5. Matrices Diagonalizables

En general, hacer aritmética con matrices es costoso, sobre todo si hay que realizar multiplicaciones entre ellas. Sin embargo, los cálculos se facilitan si las matrices son triangulares y aún más si las matrices son diagonales. De otro lado, el cálculo de los valores propios de una matriz triangular y en especial si es diagonal es trivial; por ello, sería muy conveniente encontrar un procedimiento que nos permitiera relacionar los valores propios de una matriz con los de una matriz triangular o diagonal. En el Capítulo 1, vimos un procedimiento que relacionaba una matriz con una triangular (el Método de Eliminación de Gauss), pero desafortunadamente los valores propios de la matriz triangular encontrada por este método no se relacionan con los de la matriz inicial.

Por ahora, veamos que si dos matrices son semejantes (ver Definición 5 del Capítulo 5), algunas de sus características o propiedades son iguales; entre otras, sus valores propios son los mismos y los vectores propios de una se pueden calcular a partir de los de la otra, como lo resumimos en el siguiente teorema

Teorema 7 [*Propiedades de las matrices semejantes*].
Sean A y B dos matrices semejantes, entonces

1. $\det A = \det B$.
2. Si una de las matrices es invertible, la otra también lo es.

3. A y B tienen el mismo polinomio característico.

4. A y B tienen los mismos valores propios.

5. Existe P, una matriz invertible, tal que si $\mathbf{v}$ es un vector propio de A, $P\mathbf{v}$ es un vector propio de B.

6. A y B tienen el mismo rango.

Demostración: Como A y B son semejantes, existe P, una matriz invertible, tal que $PA = BP$.

1. Puesto que $PA = BP$, $\det(PA) = \det(BP)$; es decir, $\det(P)\det(A) = \det(B)\det(P)$ y por ser P invertible, $\det P \neq 0$, lo que implica que $\det A = \det B$.

2. Por el resultado anterior, $\det A \neq 0$, si y solo si, $\det B \neq 0$, de donde concluimos que las matrices A y B son ambas invertibles o ambas no invertibles.

3. Como

$$\begin{aligned} \det(A - \lambda I) &= \det(PBP^{-1} - \lambda PP^{-1}) \\ &= \det[P(B - \lambda I)P^{-1}] \\ &= \det P \det(B - \lambda I) \det P^{-1} \\ &= \det(B - \lambda I), \end{aligned}$$

concluimos que los polinomios característicos de A y B son iguales.

4. Puesto que los polinomios característicos de A y B son iguales, sus valores propios son los mismos.

5. Sea $\mathbf{v}$ un vector propio de A asociado al valor propio λ, entonces $A\mathbf{v} = \lambda\mathbf{v}$. Multiplicando por P, tenemos $PA\mathbf{v} = \lambda P\mathbf{v}$, y por lo tanto, $BP\mathbf{v} = \lambda P\mathbf{v}$, lo cual significa que $\mathbf{z} = P\mathbf{v}$ es un vector propio de B asociado al valor propio λ.

6. La dejamos como ejercicio para el lector. □

6.6. Valores y vectores propios de una transformación lineal

Las propiedades sobre matrices semejantes establecidas en el teorema anterior y el hecho que dos matrices semejantes representen una misma transformación (ver Sección 5.3), justifican los comentarios sobre vectores propios de una transformación lineal planteados en la introducción del presente capítulo. Empecemos definiendo valor y vector propio de una transformación lineal de un espacio vectorial en si mismo.

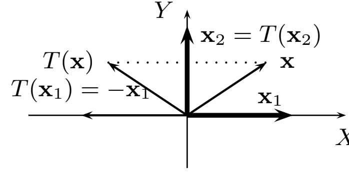
Definición 6 [*Valor y vector propio de una transformación lineal*]. Sean V, un espacio vectorial real finito, y $T : V \longrightarrow V$, una transformación lineal. Diremos que el escalar λ es un valor propio de T y que $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$, es un vector propio de T asociado a λ, si y solo si, $T(\mathbf{v}) = \lambda\mathbf{v}$ y $\lambda \in R$.

Ejemplo 16. Si $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ tal que $T(a + bx + cx^2) = (a - b) + bx + (a - c)x^2$, podemos verificar que

$$\begin{aligned} T(2 + x^2) &= 2 + x^2 \\ &= 1(2 + x^2), \end{aligned}$$

de donde concluimos que $\lambda = 1$ es un valor propio de T y que $2 + x^2$ es un vector propio de T asociado al valor propio $\lambda = 1$. □

Ejemplo 17. Determinemos geométricamente los valores y vectores propios de la reflexión en el plano a través del eje Y.



Observemos que los únicos vectores que conservan su dirección están sobre los ejes coordenados. En efecto, para los vectores sobre el eje X tenemos que $T\begin{pmatrix} a \\ 0 \end{pmatrix} = \begin{pmatrix} -a \\ 0 \end{pmatrix} = -1\begin{pmatrix} a \\ 0 \end{pmatrix}$, por tanto son vectores propios asociados al valor propio $\lambda = -1$ y para los vectores sobre el eje Y tenemos que $T\begin{pmatrix} 0 \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} = 1\begin{pmatrix} 0 \\ b \end{pmatrix}$, por tanto son vectores propios asociados al valor propio $\lambda = 1$. □

Establezcamos ahora la relación entre la definición anterior y la Definición 1 (valor y vector propio de una matriz), la cual plantea que para calcular los valores y vectores propios de una transformación lineal basta con calcular los valores y vectores propios de una matriz asociada a la transformación lineal.

Teorema 8 [*Relación entre los valores y vectores propios de una transformación y una matriz asociada*]. Sean V un espacio vectorial finito, $\mathcal{B}$ una base de V, $T : V \longrightarrow V$ una transformación lineal y A_T la matriz asociada a T en la base $\mathcal{B}$. El escalar λ es un valor propio de T y $\mathbf{v} \in V$ es un vector propio de T asociado a λ, si y solo si, el escalar λ es un valor propio de A_T y $[\mathbf{v}]_{\mathcal{B}}$ es un vector propio de A_T asociado a λ.

Demostración: Sean λ un valor propio de T y $\mathbf{v} \in V$ un vector propio de T asociado a λ. Por la Definición 6,

$$T(\mathbf{v}) \quad = \quad \lambda\mathbf{v}. \tag{6.7}$$

Como,

$$[T(\mathbf{v})]_{\mathcal{B}} = A_T[\mathbf{v}]_{\mathcal{B}} \quad \text{y} \quad [\lambda\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}} \ ,$$

reemplazando en (6.7), tenemos

$$A_T[\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}},$$

lo que significa que λ es valor propio de A_T y $[\mathbf{v}]_{\mathcal{B}}$ es un vector propio de A_T asociado a λ.

En forma similar, si λ es valor propio de A_T y $\mathbf{x} \in R^n$ es un vector propio de A_T asociado a λ, podemos concluir que λ es un valor propio de T y $\mathbf{v}$ es un vector propio de T asociado a λ, donde $\mathbf{v} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \ldots + x_n\mathbf{u}_n$, $x = (x_1, x_2, \ldots, x_n)^T$ y $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, de tal forma que $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$. □

Ejemplo 18. Si T es la transformación del Ejemplo 16, $\mathcal{B}$ es la base canónica de $\mathcal{P}_2$ y $\mathcal{B}' = \{1, 1+x, 1+x+x^2\}$, entonces la matriz asociada a T respecto a la base $\mathcal{B}$ es

$$A_T = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \qquad \text{y} \qquad [2+x^2]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

y la matriz asociada a T respecto a la base $\mathcal{B}'$ es

$$B_T = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad \text{y} \qquad [2+x^2]_{\mathcal{B}'} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}.$$

En el Ejemplo 16, vimos que $\lambda = 1$ es un valor propio de T y que $2 + x^2$ es un vector propio de T asociado al valor propio $\lambda = 1$. No es difícil verificar que

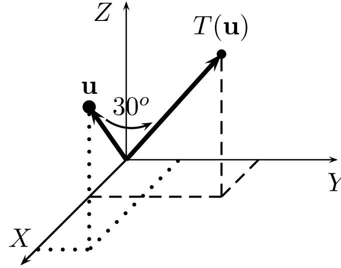
$$A_T[2+x^2]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

y que

$$B_T[2+x^2]_{\mathcal{B}'} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix},$$

de donde concluimos que $\lambda = 1$ es un valor propio tanto de A_T, como de B_T y que $[2+x^2]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ es un vector propio de A_T asociado al valor propio $\lambda = 1$ y que $[2+x^2]_{\mathcal{B}'} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ es un vector propio de B_T asociado al valor propio $\lambda = 1$. $\square$

Ejemplo 19. Determinemos los valores y vectores propios de la rotación de 30° alrededor del eje Z en R^3 geométricamente y utilizando una matriz asociada a la transformación.



Geométricamente, podemos ver que los únicos vectores que bajo la transformación conservan su dirección están en el eje Z, ya que $T\begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}$, lo que nos permite concluir que los vectores contenidos en el eje Z son vectores propios asociados al valor propio $\lambda = 1$. De otro lado, como $T(\mathbf{e}_1) = \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \\ 0 \end{pmatrix}$, $T(\mathbf{e}_2) = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \\ 0 \end{pmatrix}$ y $T(\mathbf{e}_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, la matriz asociada a la transformación respecto a la base canónica de R^3 es $A_T = \begin{pmatrix} \sqrt{3}/2 & -1/2 & 0 \\ 1/2 & \sqrt{3}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Después de una serie de cálculos, encontramos que el polinomio característico de A_T es $p(\lambda) = (1-\lambda)(4 - 2\sqrt{3}\lambda + \lambda^2)$, el cuál sólo tiene una raíz real $\lambda = 1$. Los vectores propios asociados a $\lambda = 1$ son las soluciones no triviales de $(A - I)\mathbf{x} = \mathbf{0}$, las cuales podemos verificar que son todos los múltiplos no nulos del vector $\mathbf{e}_3$; es decir, los vectores no nulos contenidos en el eje Z, como habíamos observado geométricamente. $\square$

Observemos que, al calcular los valores propios de la matriz asociada a la transformación lineal, obtenemos los valores propios de la transformación, pero que, al calcular los vectores propios de la matriz asociada a la transformación, obtenemos las coordenadas (en la base usada para calcular la matriz asociada) de los vectores propios de la transformación.

Como mencionamos en la introducción, si la matriz asociada a una transformación lineal fuese diagonal, seria muy fácil analizar la acción de la transformación; de allí la importancia de hallar la base en que la matriz asociada a la transformación sea diagonal, la cual sería fácil de calcular si dada una matriz podemos calcular una matriz diagonal semejante a ella. El que una matriz sea semejante a una diagonal, lo llamaremos de una manera especial.

Definición 7 [*Matriz diagonalizable*]. Diremos que una matriz A es *diagonalizable*, si y solo si, existe una matriz diagonal D semejante a A; es decir, que existe P, una matriz invertible y D una matriz diagonal, tal que $AP = PD$. Al procedimiento de encontrar la matriz diagonal D y la matriz invertible P lo llamamos proceso de *diagonalización* y decimos que P y D *diagonalizan* a A.

Ejemplo 20. Verifiquemos que las matrices $P = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ y $D = \begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix}$ diagonalizan la matriz $A = \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}$.

$$\begin{aligned} AP &= \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -4 & 6 \\ 4 & 6 \end{pmatrix} \\ PD &= \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} -4 & 6 \\ 4 & 6 \end{pmatrix} \end{aligned}$$

Como $AP = PD$, concluimos que P y D diagonalizan a A. □

Ahora, cómo determinar si una matriz es diagonalizable? Cuáles son las matrices que diagonalizan a A? Las respuestas a estas preguntas las establecemos en el siguiente teorema.

Teorema 9 [*Caracterización de las matrices que diagonalizan a una matriz*].
Sea A una matriz $n \times n$.

1. Si P y D son las matrices que diagonalizan a A, los elementos de la diagonal de D son los valores propios de A y las columnas de la matriz P son los vectores propios de A asociados a los valores propios, en su orden respectivo.
2. Si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es un conjunto de n vectores propios *l.i.* de A asociados a los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectivamente, entonces la matriz A se puede diagonalizar con $P = [\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_n]$ y D, la matriz diagonal tal que $d_{ii} = \lambda_i$, para $i = 1, 2, \ldots, n$.

Demostración:

1. Si
$$P = [\mathbf{p}_1\ \mathbf{p}_2\ \ldots\ \mathbf{p}_n] \quad \text{y} \quad D = \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix},$$
entonces
$$PD = [d_{11}\mathbf{p}_1\ d_{22}\mathbf{p}_2\ \ldots\ d_{nn}\mathbf{p}_n] \quad \text{y} \quad AP = [A\mathbf{p}_1\ A\mathbf{p}_2\ \ldots\ A\mathbf{p}_n].$$
Además, si P y D son las matrices que diagonalizan a A, entonces $AP = PD$. Así que $A\mathbf{p}_1 = d_{11}\mathbf{p}_1$, $A\mathbf{p}_2 = d_{22}\mathbf{p}_2$, ..., $A\mathbf{p}_n = d_{nn}\mathbf{p}_n$. De donde, concluimos que $d_{11}, d_{22}, \ldots, d_{nn}$ son los valores propios de A y que $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ son los vectores propios de A asociados a $d_{11}, d_{22}, \ldots, d_{nn}$, respectivamente.
2. Si A tiene n vectores propios $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ que forman un conjunto de vectores *l.i.*, con valores propios correspondientes $\lambda_1, \lambda_2, \ldots, \lambda_n$ y tomamos $P = [\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_n]$ y $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$, tenemos que
$$\begin{aligned} AP &= [A\mathbf{v}_1\ A\mathbf{v}_2\ \ldots\ A\mathbf{v}_n] \\ &= [\lambda_1\mathbf{v}_1\ \lambda_2\mathbf{v}_2\ \ldots\ \lambda_n\mathbf{v}_n] \\ &= PD, \end{aligned}$$
de modo que A es diagonalizable y P y D diagonalizan a A. □

Ejemplo 21. Sea $A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$. En el Ejemplo 13, vimos que los valores propios de A son $\lambda_1 = 10$ de multiplicidad algebraíca 1 y $\lambda_2 = 1$ de multiplicidad algebraíca 2 y que los conjuntos $B_1 = \left\{ \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \right\}$ y $B_2 = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \right\}$ son bases de sus respectivos espacios propios. Así que, por el Teorema 5,

$$\left\{ \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \right\}$$

es un conjunto $l.i.$ y por tanto

$$P = \begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

es una matriz invertible.

Verifiquemos que $AP = PD$, siendo D la matriz diagonal formada con los valores propios de A.

$$\begin{aligned} AP &= \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 20 & -1 & -1 \\ 20 & 1 & 0 \\ 10 & 0 & 2 \end{pmatrix} \\ PD &= \begin{pmatrix} 2 & -1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 20 & -1 & -1 \\ 20 & 1 & 0 \\ 10 & 0 & 2 \end{pmatrix}. \end{aligned}$$

Podemos concluir entonces que la matriz P de vectores propios de A y la matriz diagonal D, cuya diagonal son los valores propios de A, diagonalizan a A. □

El teorema anterior lo podemos resumir en el siguiente corolario, cuya demostración dejamos como ejercicio para el lector.

Corolario 9.1 [*Caracterización de las matrices diagonalizables*].
La matriz A es diagonalizable, si y solo si, A tiene n vectores propios que forman un conjunto $l.i.$

Ejemplo 22. Sea $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$. Puesto que $\lambda_1 = 0$ es un valor propio de A de multiplicidad algebraica 2 y $E_0 = Gen\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$, es decir, $\lambda_1 = 0$ es de multiplicidad geométrica 1, no hay un conjunto $l.i.$ formado por dos vectores propios de A, lo que implica que A no es diagonalizable. □

Teniendo en cuenta que si todos los valores propios son diferentes, entonces existen n vectores propios que forman un conjunto $l.i.$, tenemos una condición suficiente para que una matriz sea diagonalizable, la cual expresaremos en el siguiente corolario, cuya demostración también dejamos como ejercicio para el lector.

Corolario 9.2 [*Condición suficiente para la diagonalización de una matriz*].
Sea A una matriz $n \times n$. Si A tiene n valores propios diferentes, A es diagonalizable.

Ejemplo 23. Verifiquemos que $A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -1 & 3 & 0 \end{pmatrix}$ es diagonalizable. Después de una serie de cálculos, tenemos que el polinomio característico de A es $p(\lambda) = \lambda(1 - \lambda)(1 + \lambda)$, por lo tanto, los valores propios de A son $\lambda_1 = 1$, $\lambda_2 = -1$ y $\lambda_3 = 0$. Como A es de orden 3 y tiene 3 valores propios distintos, por el Corolario

9.2, la matriz A es diagonalizable. En efecto, después de una serie de cálculos, podemos determinar que los espacios propios de A son $E_1 = Gen\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}\right\}$, $E_{-1} = Gen\left\{\begin{pmatrix}0\\1\\-1\end{pmatrix}\right\}$ y $E_0 = Gen\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$, por lo tanto, $P = \begin{pmatrix}1&0&0\\1&1&0\\0&-1&1\end{pmatrix}$ y $D = \begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix}$ diagonalizan a A. Existen otras matrices P y D que diagonalicen a A? □

Debe ser claro que para que una matriz de orden n sea diagonalizable no es necesario que todos los valores propios sean diferentes, basta con que exista un conjunto $l.i.$ formado por n vectores propios. Es decir, que el polinomio característico no tenga raíces complejas y que las multiplicidades algebraica y geométrica de cada uno de los valores propios coincidan, de tal manera que en total se tengan n vectores propios que formen un conjunto $l.i.$, lo cual formalizamos en el siguiente teorema, cuya demostración es similar a la del Corolario 5.1 y la dejamos como ejercicio para el lector.

Teorema 10 [*Caracterización de las matrices diagonalizables*].
Sea A una matriz $n\times n$. La matriz A es diagonalizable, si y solo si, las multiplicidades algebraicas y geométricas de los valores propios de A son iguales y la suma de las multiplicidades algebraicas de los valores propios es n.

Ejemplo 24. Como vimos en el Ejemplo 13, la matriz $A = \begin{pmatrix}5&4&2\\4&5&2\\2&2&2\end{pmatrix}$ tiene valores propios $\lambda_1 = 10$, de multiplicidad algebraica 1, y $\lambda_2 = 1$, de multiplicidad algebraica 2, y puesto que

$$E_{10} = Gen\left\{\begin{pmatrix}2\\2\\1\end{pmatrix}\right\} \qquad \text{y} \qquad E_1 = Gen\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\2\end{pmatrix}\right\},$$

las multiplicidades geométricas de $\lambda_1 = 10$ y $\lambda_1 = 1$ son 1 y 2, respectivamente. Por el Teorema 10, la matriz A es diagonalizable. En efecto, las matrices $P = \begin{pmatrix}2&-1&-1\\2&1&0\\1&0&2\end{pmatrix}$ y $D = \begin{pmatrix}10&0&0\\0&1&0\\0&0&1\end{pmatrix}$ diagonalizan la matriz A. □

Al margen de los valores y vectores propios, la semejanza de una matriz A con una matriz diagonal permite simplificar enormemente el cálculo de las potencias de A, sobre todo si el tamaño de la matriz o la potencia son grandes. En efecto, si P y D diagonalizan a A, entonces $A = PDP^{-1}$, por lo tanto,

$$\begin{aligned} A^k &= (PDP^{-1})^k \\ &= \underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{k\ veces} \\ &= PD(P^{-1}P)D(P^{-1}P)\cdots(P^{-1}P)DP^{-1} \\ &= PDIDI\cdots IDP^{-1} \\ &= PD^kP^{-1}, \end{aligned}$$

lo cual formalizamos en el siguiente teorema y demostramos usando inducción matemática.

Teorema 11 [*Cálculo de potencias de una matriz*].
Sean P, una matriz invertible, y D, una matriz diagonal, las matrices que diagonalizan a A. Entonces

$$A^k = PD^kP^{-1}.$$

Demostración: Si P y D diagonalizan a A, entonces $A = PDP^{-1}$; es decir, el teorema se satisface para $k = 1$. Supongamos que el teorema se satisface para $k = r$; es decir, $A^r = PD^rP^{-1}$. Demostremos que se

satisface para $k = r + 1$.

$$A^{r+1} = A^r A = (PD^r P^{-1})(PDP^{-1}) = (PD^r)(P^{-1}P)(DP^{-1}) = P(D^r D)P^{-1} = PD^{r+1}P^{-1}.$$

□

Recordemos que, en la práctica, para calcular A^k es más eficiente (requiere menos operaciones) resolver los sistemas resultantes de la ecuación matricial $PA^k = D^k P$, que calcular P^{-1} y luego multiplicar.

Ejemplo 25. Sea $A = \begin{pmatrix} 3 & -4 \\ 2 & -3 \end{pmatrix}$. Verifiquemos que $A^k = I$, para todo $k = 2,\ 4,\ 6,\ \ldots$.

Después de una serie de cálculos, encontramos que $\lambda_1 = 1$ y $\lambda_2 = -1$ son los valores propios de A. Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ dos vectores propios asociados a λ_1 y λ_2, respectivamente. Por el Teorema 11, $A^k = PD^kP^{-1}$, donde $P = [\mathbf{v}_1\ \ \mathbf{v}_2]$ y $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Como $A^k = \begin{pmatrix} 1^k & 0 \\ 0 & (-1)^k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$, para $k = 2,\ 4,\ 6,\ \ldots$, tenemos que $A^k = PD^kP^{-1} = PIP^{-1} = PP^{-1} = I$, para $k = 2,\ 4,\ 6,\ \ldots$. □

6.7. Matrices Simétricas y Diagonalización

Debido a que las matrices simétricas satisfacen los resultados de diagonalización de manera especial, en esta sección, nos restringiremos a este tipo de matrices .

Para comenzar, tenemos que las raíces del polinomio característico de las matrices simétricas son todas reales, razón por la cual, las matrices simétricas tienen tantos valores propios (contando multiplicidades) como número de filas. Este resultado lo consignamos en el siguiente teorema, omitiendo su demostración por no estar al alcance de este texto[6].

Teorema 12 [*Número de valores propios de una matriz simétrica*].
Toda matriz simétrica de tamaño $n \times n$ tiene n valores propios, contando multiplicidades.

Ejemplo 26. Verifiquemos el teorema anterior para la matriz A del Ejemplo 13. A es una matriz simétrica de tamaño 3×3 y los valores propios de A son 3, contando multiplicidades: $\lambda_1 = 10$ de multiplicidad algebraica 1 y $\lambda_2 = 1$ de multiplicidad algebraica 2. □

De este teorema y el Corolario 9.2, podemos asegurar que si los valores propios de una matriz simétrica son diferentes, entonces la matriz es diagonalizable. Sin embargo, las matrices simétricas tienen más propiedades especiales.

En el Teorema 5, los vectores propios de una matriz asociados a valores propios diferentes forman un conjunto *l.i.* Para matrices simétricas, el resultado es más fuerte: los vectores propios asociados a valores propios diferentes resultan ser ortogonales, como la planteamos en el siguiente teorema.

Teorema 13 [*Ortogonalidad de los vectores propios de una matriz simétrica*].
Si A es una matriz simétrica, los vectores propios de A asociados a valores propios diferentes son ortogonales.

Demostración: Sean $\mathbf{u}$ y $\mathbf{v}$ vectores propios de A asociados a los valores propios λ y μ, respectivamente, con $\lambda \neq \mu$. Entonces,

$$\lambda\mathbf{u} = A\mathbf{u} \tag{6.8}$$

y

$$\mu\mathbf{v} = A\mathbf{v}. \tag{6.9}$$

[6] La demostración requiere de la definición del producto interno en los complejos: para $\mathbf{x}, \mathbf{y} \in C^n$,

$$\mathbf{x} \cdot \mathbf{y} = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n,$$

donde $\bar{y}$ denota el conjugado de y ($\bar{y} = \alpha - \beta i$ si $y = \alpha + \beta i$).

Multiplicando (6.8) por $\mathbf{v}^T$ y (6.9) por $\mathbf{u}^T$, tenemos

$$\begin{aligned} \lambda \mathbf{v}^T \mathbf{u} &= \mathbf{v}^T A \mathbf{u} \qquad (6.10) \\ \mu \mathbf{u}^T \mathbf{v} &= \mathbf{u}^T A \mathbf{v}. \qquad (6.11) \end{aligned}$$

Tomando la transpuesta a ambos lados de (6.11) y como $A^T = A$, tenemos

$$\mu \mathbf{v}^T \mathbf{u} = \mathbf{v}^T A^T \mathbf{u} \;\; = \;\; \mathbf{v}^T A \mathbf{u}. \qquad (6.12)$$

Restando (6.12) de (6.10), tenemos

$$(\lambda - \mu)\mathbf{v}^T \mathbf{u} = 0.$$

Como $\lambda \neq \mu$, entonces $\mathbf{v}^T \mathbf{u} = 0$, de donde concluimos que $\mathbf{u}$ y $\mathbf{v}$ son ortogonales. □

Ejemplo 27. Si tomamos la matriz $A = \begin{pmatrix} -2 & -2 & 2 \\ -2 & 4 & 8 \\ 2 & 8 & 4 \end{pmatrix}$, los espacios propios de A son

$$\begin{aligned} E_0 &= Gen\left\{\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\right\} = \left\{\begin{pmatrix} 2t \\ -t \\ t \end{pmatrix} : \; t \in R\right\} \quad , \\ E_{12} &= Gen\left\{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\} = \left\{\begin{pmatrix} 0 \\ s \\ s \end{pmatrix} : \; s \in R\right\} \quad \text{y} \\ E_{-6} &= Gen\left\{\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\right\} = \left\{\begin{pmatrix} -r \\ -r \\ r \end{pmatrix} : \; r \in R\right\} . \end{aligned}$$

Veamos que

$$\begin{pmatrix} 2t \\ -t \\ t \end{pmatrix} \cdot \begin{pmatrix} 0 \\ s \\ s \end{pmatrix} = 0\,, \quad \begin{pmatrix} 2t \\ -t \\ t \end{pmatrix} \cdot \begin{pmatrix} -r \\ -r \\ r \end{pmatrix} = 0\,, \quad \begin{pmatrix} 0 \\ s \\ s \end{pmatrix} \cdot \begin{pmatrix} -r \\ -r \\ r \end{pmatrix} = 0\,,$$

para todo $r, s, t \in R$; por lo tanto, concluimos que los vectores propios asociados al valor propio $\lambda_1 = 0$ son ortogonales a los vectores propios asociados al valor propio $\lambda_1 = 12$, que los vectores propios asociados al valor propio $\lambda_1 = 0$ son ortogonales a los vectores propios asociados al valor propio $\lambda_1 = -6$ y que los valores propios asociados al valor propio $\lambda_1 = 12$ son ortogonales a los vectores propios asociados al valor propio $\lambda_1 = -6$. □

Otra propiedad importante de las matrices simétricas, básica para la demostración del resultado principal de esta sección, es el que enunciamos en el siguiente teorema, cuya demostración es similar a la del Teorema 3 y que también omitimos.

Teorema 14 [*Igualdad de multiplicidades de los valores propios de una matriz simétrica*].
Si A es una matriz simétrica, las multiplicidades algebraica y geométrica de cada uno de los valores propios de A son iguales.

Ejemplo 28. Volviendo sobre la matriz del Ejemplo 13, $A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$, tenemos que A tiene valores propios $\lambda_1 = 10$ y $\lambda_1 = 1$, de multiplicidades algebraica 1 y 2 y multiplicidades geométricas 1 y 2, respectivamente. □

El que una matriz sea diagonalizable, como nos dice el Teorema 9, significa que existe una base de vectores propios que son las columnas de la matriz P que la diagonaliza. En algunos casos especiales, las columnas

de la matriz P las podemos escoger *ortonormales* o dicho de otra forma, la matriz P resulta ser ortogonal. Esta particularidad de la matriz P nos conduce a la siguiente definición.

Definición 8 [*Matriz Ortogonalmente Diagonalizable*]. Decimos que una matriz A es *ortogonalmente diagonalizable*, si y solo si, existe una matriz diagonal D tal que $AQ = QD$, donde Q es una matriz ortogonal. En otras palabras, A es *ortogonalmente diagonalizable*, si y solo si, A es diagonalizable mediante una matriz ortogonal.

Ejemplo 29. Si A es la matriz del Ejemplo 27 y $Q = \begin{pmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}$, no es difícil verificar que $AQ = QD$, siendo D la matriz diagonal formada con los vectores propios de A, de donde concluimos que A es ortogonalmente diagonalizable . □

Del Teorema 14 y el Corolario 5.1, tenemos que si A es una matriz simétrica $n \times n$, existe una base $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ de R^n, donde $\mathcal{B}_i$ es una base del espacio propio correspondiente al valor propio λ_i. Podemos suponer que cada una de las bases $\mathcal{B}_i$ es ortonormal (si no lo son, con el Proceso de Grand-Schmidt las podemos ortonormalizar). Puesto que el Teorema 12 nos dice que los vectores propios asociados a valores propios diferentes de una matriz simétrica son ortogonales, concluimos que podemos construir a $\mathcal{B}$ como una base ortonormal de vectores propios de A. En otras palabras, tomando la matriz Q como la matriz de los vectores de $\mathcal{B}$, podemos concluir que la matriz A es diagonalizable por Q, demostrando una de las implicaciones del principal resultado de esta sección, expresado en el siguiente teorema.

Teorema 15 [*Caracterización de una matriz simétrica*].
La matriz A es simétrica, si y solo si, es ortogonalmente diagonalizable.

Demostración: Nos queda por demostrar únicamente que si A es ortogonalmente diagonalizable, entonces A es simétrica. Que A sea ortogonalmente diagonalizable significa que

$$A = QDQ^T, \tag{6.13}$$

donde Q es una matriz ortogonal (recordemos que si Q es ortogonal $Q^{-1} = Q^T$). Por lo tanto,

$$A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A,$$

de donde concluimos que A es simétrica.

Ejemplo 30. Volviendo sobre la matriz del Ejemplo 27, tenemos que $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}$ es una base ortogonal de R^3, conformada por vectores propios de A. Al normalizar cada uno de los vectores de $\mathcal{B}$, obtenemos una base ortonormal de R^3, $\mathcal{B}' = \left\{ \begin{pmatrix} 2/\sqrt{6} \\ -1/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\}$. Si tomamos $Q = \begin{pmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}$, como vimos en el ejemplo anterior, Q y la matriz diagonal formada con los valores propios de A diagonalizan a A.

6.8. Ejercicios

1. Determine cuales de los vectores dados son vectores propios de la matriz A dada y en caso de serlo, identifique los valores propios asociados.

   a) $A = \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix}$, $\mathbf{v}_1 = \begin{pmatrix} -3 \\ 3 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 2 \\ -4 \end{pmatrix}$.

b) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 2 & 1 & 2 \end{pmatrix}$, $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 9 \\ 0 \\ -6 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -6 \\ 4 \end{pmatrix}$.

c) $A = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 1 & 2 \\ -1 & -2 & 0 \end{pmatrix}$, $\mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} 5 \\ 2 \\ -3 \end{pmatrix}$.

2. Determine cuales de los escalares dados son valores propios para la matriz A dada.

   a) $A = \begin{pmatrix} 0 & 9 \\ 4 & 0 \end{pmatrix}$, $\lambda_1 = 0$, $\lambda_2 = 6$, $\lambda_3 = 9$, $\lambda_4 = -6$.

   b) $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 3 & 0 & 2 \end{pmatrix}$, $\lambda_1 = 0$, $\lambda_2 = 4$, $\lambda_3 = -1$, $\lambda_4 = 2$.

   c) $A = \begin{pmatrix} 1/2 & 2 & 3 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 3 & 2/5 \\ 0 & 0 & 0 & 2/5 \end{pmatrix}$, $\lambda_1 = 0$, $\lambda_2 = 2$, $\lambda_3 = 1/2$, $\lambda_4 = 2/5$.

3. Cuáles son las multiplicidades algebraicas de los valores propios de la matriz del Ejemplo 2?

4. Determine los valores propios, sus multiplicidades algebraicas y sus multiplicidades geométricas, para cada una de las matrices A dadas.

   a) $A = \begin{pmatrix} 0 & -9 \\ 4 & 0 \end{pmatrix}$

   b) $A = \begin{pmatrix} -2 & 2 & 3 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

   c) $A = \begin{pmatrix} 3 & 0 & 3 \\ 0 & -1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$.

   d) $A = \begin{pmatrix} 2 & 0 & 3 & 1 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

5. Determine, para cada una de las matrices del Ejercicio 3, en qué casos existe una base de R^n, conformada por vectores propios de la matriz A. En caso de existir, exhíbala.

6. Demuestre el Corolario 5.1.

7. Demuestre el Teorema 6.

8. Conteste falso o verdadero a cada una de las siguientes afirmaciones, justificando su respuesta.

   a) Los valores propios de una matriz son raíces de su polinomio característico.

   b) Las raíces del polinomio característico de una matriz son sus valores propios.

   c) Si 0 es un valor propio de una matriz A, entonces A es la matriz nula.

   d) Toda matriz $n \times n$ tiene n valores propios, contando sus multiplicidades algebraicas.

   e) Todos los valores propios de una matriz nula son 0.

   f) Cualquier vector de R^n es vector propio de la matriz nula $n \times n$.

   g) Si dos matrices tienen los mismos valores propios, entonces sus vectores propios son los mismos.

   h) Para que exista una base de R^n, conformada por vectores propios de una matriz $n \times n$, todos sus valores propios deben ser diferentes.

   i) Para conocer los valores propios de la matriz inversa de A no es necesario encontrar la inversa.

   j) La matriz $5I_n$ tiene sus n valores propios iguales a 5.

9. En cada caso, con la información dada sobre la matriz A de tamaño $n \times n$, determine si A es diagonalizable, ortogonalmente diagonalizable, invertible y/o simétrica y el valor de n.

   a) El polinomio característico de A es $p(\lambda) = (\lambda - 2)(\lambda + 3)(\lambda + \sqrt{2})\lambda$.

   b) El polinomio característico de A es $p(x) = (x - 1)(x + 1)(x^2 + 1)$.

   c) El polinomio característico de A es $p(\alpha) = (\alpha - 4)^2(\alpha + 2)^3$.

   d) Los espacios propios de A son $E_2 = Gen\left\{\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}$ y $E_{-1} = Gen\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}$.

   e) Los espacios propios de A son $E_3 = Gen\left\{\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}\right\}$ y $E_0 = Gen\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}\right\}$.

   f) $A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 0 & 3 \\ 2 & 3 & -5 \end{pmatrix}$.

10. Demuestre que si A es una matriz diagonalizable cuyos valores propios son 1 y -1, entonces A es invertible y $A^{-1} = A$. Además, $A^k = \begin{cases} A & \text{si } k \text{ es impar} \\ I & \text{si } k \text{ es par} \end{cases}$. (AYUDA: Ver Ejemplo 25)

11. Conteste falso o verdadero, justificando su respuesta.

   a) Si una matriz 5×5 tiene 5 valores propios diferentes, es diagonalizable.

   b) Si $T : R^4 \longrightarrow R^4$ es una transformación lineal y A es su matriz asociada en la base usual, los valores y vectores propios de A son los valores y vectores propios de T.

   c) Si $T : R^6 \longrightarrow R^6$ es una transformación lineal y A es su matriz asociada en una base diferente a la usual, los valores y vectores propios de A son los valores y vectores propios de T.

   d) Si $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ es una transformación lineal , T no tiene valores ni vectores propios.

   e) Si $T : \mathcal{M}_{2\times2} \longrightarrow \mathcal{P}_4$ es una transformación lineal , T no tiene valores ni vectores propios.

   f) Si A es simétrica, A es diagonalizable.

   g) Si $A = A^T$ es una matriz de tamaño 5×5, A tiene 5 valores propios diferentes.

   h) Si $A = A^T$ es una matriz de tamaño 9×9, A tiene 9 vectores propios que forman un conjunto ortogonal.

   i) Si $A = A^T$ es una matriz de tamaño 7×7, cualquier conjunto de 7 vectores propios es ortogonal.

   j) Las matrices que diagonalizan una matriz dada son únicas.

Índice alfabético